POWER CONTROL AND CAPACITY ANALYSIS IN COGNITIVE RADIO NETWORKS

A Dissertation
Presented to
The Academic Faculty

By

Pan Zhou

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in
Electrical and Computer Engineering

School of Electrical and Computer Engineering
Georgia Institute of Technology
August 2011

POWER CONTROL AND CAPACITY ANALYSIS IN COGNITIVE RADIO NETWORKS

Approved by:

Dr. John A. Copeland, Advisor
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Jeff S. Shamma
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Raheem A. Beyah
Adjunct Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Yuri Bakhtin
Associate Professor, School of Mathematics
Georgia Institute of Technology

Dr. Gordon L. Stuber
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: May 10, 2011

To Chunli Pan and Bengui Zhou – my mother and father.

ACKNOWLEDGMENTS

First of all, I am most grateful to my research advisor, Professor John A. Copeland. I do believe that he is a real mentor in my research and life. I have benefited a lot from his impressive insights in research. He also helped me improve my academic writing and oral presentation significantly. During the years when I have being in communication systems center (CSC) lab, Professor John A. Copeland's insightful advice and wisdom in both research and life influenced me a lot, and he sets a lifelong model for me to become a nice and strong person. I would like to thank him for his support, encouragement, and invaluable advice during the course of my Ph.D. study. I would also like to thank the Associate Chair of Graduate Affairs Prof. Bonnie H. Ferri. As a perfect lady, she helped me a lot for the hardness time I faced during my Ph.D. program. I am also thankful to my committee members, Prof. Raheem A. Beyah, Prof. Gordon L. Stuber, Prof. Jeff S. Shamma, and Prof. Yuri Bakhtin for their comments on my thesis. Their comments improved my dissertation.

I would like to thank my lab members, Dr. Yusun, Dr. Selcuk, Fisal, Moazzam, Sungjin and Shi-In etc. in the CSC lab for their friendship and encouragement on me.

I would also like to thank the Chinese Academy of Engineering, Prof. Peigeng Li, the president of my previous university, Huazhong University of Science and Technology (HUST), for his love, trust, encouragement and expectation on me as one of the excellent HUSTers. I am thankful for my previous lab mentor Prof. Wei Liu in my master study. He gave me an initial and important guidance in my research. I thank all my friends at Georgia Institute of Technology. You make this place vivid, warm, happy, and attractive.

Finally, I would like to thank my mother and father, Chunli Pan and Bengui Zhou. As a full professor and official in a Chinese university, they have encouraged me a lot to get my Ph.D. Especially, my mother has been educating me to work hard but work intelligently to overcome hardship since my childhood. It is she who has shaped my positive attitude in both work and life. This dissertation is dedicated to my parents.

TABLE OF CONTENTS

LIST OF FIGURES

SUMMARY

In recent years, *cognitive radio* (CR) as a novel communication paradigm has triggered great interest in the research community (see two comprehensive surveys [1, 2]). This communication paradigm is capable of dynamically sensing and accessing unused or under-utilized spectrum bands in a target licensed spectrum pool and of communicating via these spectrum bands without causing any harmful interference to the primary (licensed) users (PUs).

As indicated in the Federal Communications Commission (FCC) report [3], 90% of the licensed spectrum bands remain idle at a given time and location, while remaining spectra for further licensing become more and more scarce. To solve this problem, CR, with its ability of opportunistic and dynamic spectrum access [2], is regarded as a promising technology to effectively address the spectrum-insufficiency problem for future wireless communications. For CR users or secondary users (SUs)[1], two spectrum-access schemes exist: namely, *spectrum underlay* and *spectrum overlay* [4]. The spectrum overlay improves spectrum utilization by granting SUs the authority to sense and explore the unused spectrum offered by PUs. In this scenario, SUs can use the unused channels of PUs and transmit with full power on these channels. The spectrum underlay is another spectrum-access scheme that further improves the under-utilized spectrum resource. It permits SUs to share the same spectrum bands with PUs at the same time and location. In the spectrum overlay, designing effective and efficient *spectrum-sensing* [2] techniques in the physical layer for *cognitive radio networks* (CRNs) is the major concern. The SUs first detect the existence of active PUs on the PUs' channels; if the PU on a specific channel is detected to be idle (or inactive), then, the SUs can use this channel for their own data communications, and vice versa. In the spectrum underlay, SUs need to restrict their transmit power to not cause severe interference to PUs. So, the *power-control* problem is the major concern in

[1] the terms "cognitive radio users" and "secondary users" are used alternatively in this research.

spectrum-underlay CRNs, and it is one of the kernel problems in the resource-allocation framework [5] of CRNs.

In this research, we mainly focus on the power-control problem in the spectrum-underlay CRNs. Compared to the traditional *network-utility-maximization* (NUM) framework [6] that is used to find the optimal transmit powers for the aggregated network-throughput maximization in traditional wireless data networks such as cellular networks, ad hoc networks, and wireless sensor networks, the sharp difference from CRNs lie in the feature of opportunistic and competitive spectrum access among SUs, where the non-cooperative game theory, as a standard mathematical tool, is used to control each SU's transmit power results in each SU's own utility maximization (see a survey paper [7]). Regardless of the specific problems and/or application scenarios, this research aims to both mitigate the interference from SUs to PUs that guarantees the quality-of-service (QoS) of primary communications and to maximize the utility (or revenue) of each individual SU among all the competitive SUs in CRNs.

Then, after finishing the investigation of the power-control problem for practical CRNs, a theoretic research to understand the fundamental throughput capacity for the newly arising *cognitive radio ad hoc networks* (CRAHNs) is then conducted. This work corresponds to *the scaling law of cognitive radio ad hoc networks (CRAHNs)*. CRAHNs, as surveyed in [5], will be a type of powerful distributed wireless networks for military and civilian communications in the upcoming decades. The scaling law study of CRAHNs has gained much interest in the research community. This research is stimulated by the seminal work of Gupta and Kumar [8] on the fundamental throughput capacity scaling law for large-scale wireless ad hoc networks. It has became an active research topic in recent years, e.g., [8] [9] [10] [11] [12]. As shown in [8], the per-node[2] throughput capacity for the wireless ad hoc networks is $\Theta(W/\sqrt{n \log n})$ [3] *bits/sec* in the random and static deployment

[2] the term "node" and "user" are used interchangeably in this research.

[3] We use the same notations $g(n) = \Theta(f(n))$, $g(n) = \Omega(f(n))$, $g(n) = \omega(f(n))$, $g(n) = O(f(n))$, $g(n) = o(f(n))$, etc. as in [8]. Notation W stands for the available network bandwidths; n denotes the density of the ad hoc networks. In the following discussion, n denotes the density of CRAHNs; m denotes

condition, where W stands for the available network bandwidths and n denotes the density of the ad hoc networks. This result shows that the per-node throughput for wireless ad hoc networks decreases with the increase of the network density. It indicates that the capacity scaling of wireless ad hoc networks is not so promising when compared to wireline networks, where the constant capacity scaling $\Theta(1)$ can always be achieved. Motivated by the research on classic wireless ad hoc networks [8], the scaling-law research on CRAHNs has also triggered great interest recently (see literature review in section 1.3).

In this work, a research on CRAHNs that takes into account the impact of *PU activity* [5] is studied. "CRAHNs under the impact of *PU activity*" is a typical and important network scenario for CRAHNs that has never been considered in the research community yet. This research aims to give an in-depth study of the scaling law of CRAHNs in this network scenario. This work is believed to have unique value, and it will have an impact to the research community.

the density of primary networks.

CHAPTER 1

INTRODUCTION

1.1 Motivation

1.1.1 Power Control in Cognitive Radio Networks

In Cognitive Radio (CR) [2] networks, power control deals with the selection of proper transmit power for CR users' transmissions that achieves high spectrum efficiency by enabling CR users to reuse the PUs' spectrum bands under the interference constraints imposed by PUs. In the next generation wireless communications, CR users are expected to be uncoordinated opportunistic users, whereas there are conflicting interests among the CR users [13], [14]. This motivates the use of noncooperative game theory to do research on CR networks (see a survey paper [13]). Compared to traditional centralized solutions, the game-theoretical approach has the advantage of distributed implementation for CR networks: each CR user only takes care of their own utility maximization, and it does not need to know other users' payoff (or utility function).

1.1.1.1 Power-Control Game in Wideband Cognitive Radio Networks

The power-control problem is especially important for wideband CRNs, since the networks are *interference-limited*. As indicated in FCC report [15], CDMA is one of the natural underlay communication schemes in CRNs. Especially, for competitive CR users in CRNS, the power-control problem by applying non-cooperative game theory has been thoroughly investigated in the research community (see literature review in subsection 1.2.1).

The major contributions of the related research on power control can be summarized as follows: to find the optimal transmit power that maximizes each SU's own utility and/or guarantees the fairness among SUs. And, usually, the SUs' SINRs are used as their QoS performance metric. As we know, besides acquiring enough transmit power to guarantee reliable communication links, different SUs need to support heterogenous services with

different transmission rates. For example, some SUs require real-time multimedia transmissions, while other SUs require non-real-time data transmissions. In the wideband communication (e.g. CDMA, UWB) scenarios, a SU with a high SINR might not be allocated an adequate transmission rate to accomplish its communication task. So, to provide flexible transmission rates for SUs, efficient use of radio resources requires a transmission-rate control in addition to transmit-power control. And, this motivate our first investigation on joint power and rate control for wideband CR users, which is proved to be an effective way to improve the users' utility and network performance.

1.1.1.2 Power-Control Game in Fading and Multi-Channel Cognitive Radio Networks

To date, power control of CR users in non-fading environments gains intensively research in the research community (see literature review in subsection 1.2.1 and subsection 1.2.2). In all of these works, the authors considered quasi-stationary wireless channel model, where the channel gains are known and stationary over time. However, these assumptions are only a special case in practice, and terminal users experience frequent channel changes over time dynamically. Especially in CR networks, CR users need to update power levels frequently to maintain their SINR level (usually referred as QoS) due to channel fading. Furthermore, the transmissions outage of PUs due to channel fading and interferences from CR users should be addressed when the power control of CR users is concerned. To the best of our knowledge, existing works on power control for CR users in fading channels are only [16] and [17], in which the authors studied the ergodic capacity maxi- mization problems for a single user case. The multi-user power-control problems are of much more interest and importance in practical CR networks. However, there is no existing work to study this problem. Moreover, power control in *Orthogonal Frequency Division Multiplexing Access (OFDMA)/multi-channel* CRNs has also received much attention in the research community. Orthogonal Frequency Division Multiplexing (OFDM) is regarded as a potential transmission technology for the future CR systems due to its flexibility in allocating resources and high achievable data rate. So, motivated by the growing interest of general

multi-channel case (contains the single-channel case) performance in fading environments, we consider the multiuser power-control problem for CR networks over multiple channels using a game-theoretical approach. The fundamental performance traits of multiple CR user power control are investigated in the presence of multiple primary users (PUs) in fading channels for the first time.

1.1.1.3 Learning-based Power-Control Game Cognitive Radio Networks

As we know, game-theoretical approach provides distributive solutions for CRNs. However, to conduct the power-control game, a selection of the optimal power strategy for each CR user requires its rivals' (i.e. other CR users in the network) interference channel and transmit power strategy information. There are overhead costs for channel estimations and information exchange of power strategies among difference CR users. The information exchange among CR users by using non-cooperative game theory would scale in a order of $\Theta(N^2)$, which is infeasible in the large network deployment.

In the power-control game of CR networks, two fundamentally conflicting objectives exist in CR networks. On the one hand, each CR user wishes to achieve higher Signal-to-Interference plus Noise Ratio (SINR) and get higher transmission rates, which results in better Quality of Service (QoS); on the other hand, this higher SINR is achieved at the expense of higher interference to other CR users and to PUs. This selfish behavior of users would lead to network performance degradation [18], which is called the "*tragedy of commons*" in economics. To solve this problem, most of the existing works introduce "pricing" [19] as an incentive scheme to facilitate more efficient resource utilization for selfish CR users [20], [21], [22], [23], [24], [25]. However, the pricing-based scheme requires a certain degree of cooperation among CR users, and it is difficult to configure and implement pricing schemes in CR networks, since there lacks a centralized coordinator and/or there is no single *creditable* user in CR networks. Moreover, price-based incentive schemes can incur significant overheads in the algorithm design. For example, in [26], a self-incentive pricing scheme is adopted, where each CR user checks the best pricing

factors for every power control strategy at the Nash equilibrium. This scheme has high implementation complexity, which might not be suitable in the large network deployment.

As we have noticed, the above works [20], [21], [22], [23], [24], [25] focused on power control in the slow channel-varying or no-fading environments, where the channel information exchange need not to be performed so frequently. Power control for CR networks in fading environments has also being studied [27] [28] [29]. These works indicate that the acquisition of adequate Channel State Information (CSI) from PUs that results in better throughput for CR users in fast channel-varying environments is a hard problem.

Moreover, due to the opportunistically and heterogenous nature of CR users, CR users are most likely to be autonomous users. The usual assumption of spontaneous willingness to exchange their private information, e.g., channel and strategy information, is unrealistic for conflicting-interest CR users. Moreover, the existing conflicting interest among CR users might cause dishonest behavior in resource competition, e.g., CR users may even exchange false private information about their channel conditions in order to get more access to the spectrum [30]; some of the CR users may even emulate the PUs' behavior to attack other CR users with known channel statistics [31] and with unknown channel statistics [32].

Our proposed work is motivated by the drawbacks of classic game-theoretical approaches stated above, and the growing interest from the research community that concerns on fully distributed power-control algorithms for CR networks (see literature review in subsection 1.2.3): the demand of robust distributed power-control algorithms with low communication overheads for opportunistic and competitive CR users.

1.1.2 Capacity Analysis in cognitive radio ad hoc networks (CRAHNs)

The CRAHNs, as surveyed in [5], will be a type of powerful distributed wireless networks for military and civil communications in the next decades. The scaling law study of CRAHNs has gained much interest in the research community.

Stimulated by the seminal work of Gupta and Kumar [8], the fundamental throughput

scaling law for large-scale wireless ad hoc networks has become an active research topic [8] [9] [10] [31-37] [88-90] in recent years. It is shown in [8] that the per-node throughput capacity for the wireless ad hoc networks is $\Theta(W/\sqrt{n\log n})$[1] *bits/sec* in the random and static deployment condition. When allowing nodes to move independently and uniformly in the network region, the authors in [9] shows that a constant per-node throughput capacity can be achieved. Since throughput and delay are important network performance metrics, consequently significant effort has been devoted to research their trade-off in wireless networks [31-35]. In [10], the optimal delay-throughput tradeoff is established for both static and mobile wireless networks. The delay and throughput have also been studied under different mobility models, such as [33] [34].

Compared to classical wireless ad hoc networks, the key distinguishing factors of the dynamic network topology, and the time and the location varying spectrum availability make the analysis of scaling laws for CRAHNs very challenging. The research on scaling laws, i.e., the asymptotic throughput capacity and delay, etc., for CRAHNs is still in the infancy phase [35] [36] [37]. In [35] [36], the throughput scaling is considered for the CRAHNs under the spectrum sharing scheme of *spectrum underlay* [2] by making use of *spatial domain*, which enables secondary (unlicensed) users to coexist with PUs in the licensed spectrum bands while causing no excessive interference to PUs. It is shown that both CRAHN and primary network can simultaneously achieve the same throughput scaling law as a stand-alone network: specifically, the per-node throughput capacity of primary network and the CRAHN are $\Theta(W/\sqrt{m\log m})$ *bits/sec* and $\Theta(W/\sqrt{n\log n})$ *bits/sec* respectively. In [37], the throughput capacity of multi-channel wireless networks is investigated in the presence of constraints on channel switching. Indicated in [37], the network scenario is applicable to CRAHNs where the secondary users utilize a portion of the spectrum that is not being used by PUs (*multi-spectrum feature*). They introduce an assignment model

[1]We use the same notations $g(n) = \Theta(f(n))$, $g(n) = \Omega(f(n))$, $g(n) = \omega(f(n))$, $g(n) = O(f(n))$, $g(n) = o(f(n))$, etc. as in [1]. Notation W stands for the available network bandwidths; n denotes the density of CRAHNs; in the following discussion, m denotes the density of primary networks.

wherein a node can only switch between a set of f contiguous channels ($2 \leq f \leq c$), where c is the number of total subchannels. The per-flow capacity for the adjacent (c, f) channel assignment model is shown to be $\Theta(W\sqrt{\frac{f}{cn\log n}})$, where $f/c \leq 1$. As shown in [38], this spectrum sharing scheme is referred as *spectrum overlay* [2] by making use of the *frequency domain*. Moreover, beside the spectrum access schemes by making use of spatial domain and frequency domain, the spectrum overlay can also make use of *time domain* where secondary users utilize the time slots based on the on-and-off activity of PUs' traffic load [38] (*time-varying feature*).

To our best knowledge, the scaling laws which consider the important *time-varying feature* of spectrum sharing schemes in CRAHNs due to *PU activity* in spectrum overlay is still not touched in the research community to date. As pointed out in [5], the PU activity is the one of the dominant features that affects the secondary spectrum access, and nearly all the challenging networking issues raised in CRAHNs, such as routing, MAC, topology control and spectrum decision, etc. are its direct outcome. Moreover, its importance attracts the recent intense research on PU activity measuring and modeling [5] [39], accurate modulation of which impacts on the performance analysis and network design of CRAHNs.

1.2 Literature Review:Power Control in Cognitive Radio Networks

Cognitive radio cooperated with the advanced physical-layer communications schemes such as, code division multiple access (CDMA) and ultra-wideband (UWB) in wideband communications systems, and multi-channel/OFDMA in narrowband communications systems, generate different types of CRNs. Although they have their own unique features and advantages that further improve the spectra utilization for next generation wireless communications (see their applications in [16]), these CRNs have recently gained insensitive studies. Moreover, besides the above CRNs with the advanced physical-layer communications schemes. the growing interest concerning with more distributed and robust solutions of power-control algorithms in CRNs from the research community intrigues us to design the

power-control algorithm with low implementation-complexity for realistic CRNs. In this section, the state-of-art research progress of power control in CRNs is reviewed, which includes power control in wideband CRNs, power control in fading and mutli-channel CRNs and learning-based power control in CRNs.

1.2.1 Power-Control Game in Wideband Cognitive Radio Networks

The power-control problem for traditional cellular wireless networks was addressed in several classic papers [40–42]. The power-control problem in wireless networks with competitive users was also investigated in the literature. A non-cooperative uplink power-control game was formulated in [42], the outcome of which resulted in a Nash equilibrium that was inefficient. Therefore, a pricing scheme was introduced in order to obtain Pareto improvement.

In CRNs, two game-theoretic power-allocation schemes were proposed in [43] to achieve efficiency and fairness for SUs. Price was used as a control parameter in the CRN to achieve objectives such as efficiency. However, the utility maximization for each SU was not considered. In [44], a general power-control model was proposed for CRNs. The primary network was considered to regularly monitor the interference from all SUs, and the behavior of SUs was modeled as a non-cooperative power-control game. In that game, the utility function of each SU was the logarithm of the user's *Signal to Interference and Noise Ratio* (SINR) which represented the throughput of communication systems in the high SINR regime. The existence, uniqueness and Pareto efficiency of the Nash equilibrium were also investigated. In [20], the authors addressed the problem of utility-maximization for PUs and presented a Stackelberg game[2] model for PUs and SUs. However, it ignored the constraints on the PUs' resource usage and performance guarantees for the SUs.

The *pricing* schemes in CRNs have been considered in recent literatures. In [21], the authors investigated several different secondary pricing by capturing the effects of

[2]It is a dynamic model of duopoly proposed by Stackerlberg (1934) in which a dominant (or leader) firm moves first, and a subordinate (or follower) moves second [45].

network-wide interferences. The dynamics of price competition among competitive spectrum providers was analyzed in [22]. The problem of pricing-based utility-maximization for the spectrum owners was addressed in [23, 46]; however, the resource constraints and performance guarantees for SUs were not considered. In [25], the pricing issue was studied in a competitive CRN in which the SUs strategically adjusted their uplink transmit-power levels to maximize their own utilities, and the primary service provider charged the SUs on their transmit-power levels to enhance its own utility.

The problem of power-control and rate-adaptation for SUs in a CDMA environment was recently considered in [47]. This work adopted the classic NUM [6] framework for the SUs given the interference constraints from PUs. The SUs have a QoS constraint in terms of minimum SINR and transmission rate. As we have explained in Chapter I, this framework is not suitable for the competitive SUs in CRNs. Usually, different SUs have their own transmission rate and utility; an aggregated network throughput will not be appropriate. Thus, the behavior and equilibrium points among these competitive SUs with different utilities should be studied and understood. Moreover, a pricing scheme was not introduced in [47].

1.2.2 Power-Control Game in Fading and Multi-channel Cognitive Radio Networks

1.2.2.1 Power-Control Game in fading-channel Cognitive Radio Networks

In [16], the authors considered a CR system in fading wireless channels and proposed an opportunistic power-control strategy for a CR user, which protected the PU's transmission and realized spectrum sharing between the PU and the CR user. Via opportunistically adapting its transmit power, the CR user could maximize its achievable transmission rate without degrading the outage probability of the PU. In [48], the authors studied the optimal power-control policies for fading channels in CRNs by considering both the transmit- and the interference-power constraints. For each of the constraints, the peak power and the average power were investigated. The optimal power-allocation strategies in terms of maximizing the ergodic capacity of the SU was derived when the channel state information

was available to the transmitter and the receiver. In [49], an adaptive-power-control scheme was proposed for a cognitive radio system (CRS) in a Rayleigh-fading channel. By allowing transmit-power adaptation at the SU transmitter to maintain a constant output *Signal to Noise Ratio* (SNR) to the SU receiver, this scheme maximized the output SNR and limited the interference to a PU within an interference constraint. In [50], the authors considered a CRN where a SU shared the same narrow band with a PU for transmission. A new type of constraint for the SU to protect the primary transmission was proposed, which limited the maximum outage probability of the primary transmission subject to the SU's interference to be below a prescribed target. The optimal power-allocation strategies over block-fading channels were proposed for the SU to achieve its outage capacity.

The works [20]-[23] mainly focused on the information-theoretical capacity analysis of a single SU in a single fading-channel scenario, whereas the important network scenario of *multi-channel* and *multi-user* power control for practical CRNs in fading environments is missing.

1.2.2.2 Power-Control Game in multi-channel Cognitive Radio Networks

In a wireless network where both the primary system and the secondary system employ the OFDM (or multi-channel) transmission technology, the SUs can flexibly fill in the spectral gaps left by the PUs [51] or transmit over the unused subchannels left in the primary system [52]. Even If there are no unused subchannels left in the primary system, SU can flexibly share the subchannels with the PUs, with the interference-power constraints to protect the PUs [53]. In [53], when the SU and the PU coexist in the same bands, with individual interference-power constraint imposed on each subcarrier to protect the primary transmission, the optimal power-allocation strategy to maximize the rate of the SU was derived. In [54], the authors considered a spectrum-sharing environment where an OFDM-based CR system sharing the same band with an OFDMA primary system. A new type of constraint referred to as the rate-loss-constraint, in the form of an upper bound on the rate loss of PU due to the secondary transmission, was used to protect the PU. The optimal

power-allocation strategy was proposed to maximize the rate of SUs subject to the PUs' rate-loss constraint. In [30], the distributed multi-channel power allocation for spectrum-sharing CRNs with QoS-guarantee is studied by using the non-cooperative game theory. The problem is formulated as a non-cooperative game with coupled constraints of the co-channel interference among SUs and the interference-temperature regulation imposed by primary systems. The properties of the Nash equilibrium for the proposed game including the existence and QoS provisioning are thoroughly investigated. A layered structure by applying the Lagrangian dual-decomposition is derived, and a distributed algorithm to find the Nash equilibrium is designed via this structure.

The above work [51-54] focused on maximizing the aggregated network throughout for the SUs in a OFDMA/multi-channel-based CRNs, while work [30] aimed to study individual throughput maximization for competitive SUs in a multi-channel CRNs. To the best of our knowledge, the performance of multiple SUs in a fading multi-channel environment is not studied yet. However, this is a practical network scenario where the OFDMA-based CRNs [16] lie in the fading environments.

1.2.3 Learning-based Power-Control Game in Cognitive Radio Networks

As discussed in subsections 2.1.1 and 2.1.2, the CR users are considered to be selfish terminal users, they compete the spectrum resources of PUs to maximize their own utilities. This motivates the applying of non-cooperative game theory to do the research in CRNs, and the above preliminary research on *wideband CDMA* and *fading and multi-channel* CRNs are the representative ones.

As we know, power-control problem is one of the key research issues in spectrum underlay CR networks. Especially, the non-cooperative behaviors among CR users by applying game theory has gained intensively study in recent years in CR networks, e.g. [22], [14], [27], [20], [21], [23], [25], [26], [24]. Almost all of them assume the CSI among CR users and from CR users to PUs is known[3]. As we know, the channel estimation among CR users

[3]Usually, researchers assume the CSI estimation among the CR-Tx and CR-Rx is via classical channel

requires cooperation of all CR users, which has high communication costs. Moreover, it requires the *willingness* and *honest* behaviors of all CR users, which might not be applicable for autonomous CR users. Also, the CSI estimation from CR users to PUs is a hard problem [29][4] that researchers usually ignored, where obvious cooperation among PUs and CR users are not permitted in CR.

The *pricing* scheme in CRNs has been considered in recent literatures [20], [21], [23], [25], [26], [24], [22]. Pricing has the advantage to improve the spectrum utilization. It is also a good approach to control the interference from CR users to PUs [26] and facilitate spectrum auction for CR users offered by PUs [43]. However, it requires cooperative behaviors among CR users and/or PUs, and the implementation complexity is high.

Nowadays, the research community shows more and more interests on designing fully distributed power-control algorithms for CR networks. This is especially important for competitive and autonomous CR users for future wireless communications in tough radio-access environments. Two of the key research issues are as follows: a) how to reduce or eliminate the overhand of CSI and power strategy information exchange among CR users that grants a robust distributed network; b) how to effectively estimate and explore the CSI from PUs for CR users' power control. For the research issue b), there already have several works; there is seldom work on the research issue a). In [29], the authors proposed a novel power control scheme to maximize the capacity of a *single* CR user by exploring the PU's CSI, which is based on the measurement of the average interference-power constraint at the PU-Rx. In [55], the authors considered the scenario on utilizing the PU's ACK/NAK information to maximize the utility of a *single* CR user. However, both [55] and [29] considered only a single SU. In [56], the authors considered the power control of multiple CR users by letting them listen to the PU's feedback channel as an external inference signal for

training, estimation, and feedback mechanisms, while, for CSI estimation from the PU-Tx to the CR-Rx , they assume CR users have the preknowledge on the PR-Tx power level and the channel reciprocity, when estimating the received signal power from the PR-Tx.

[4]The PUs' CSI to the CR-Tx can be obtained by applying, 1) eavesdropping the CSI feedback from the PU-Rx to the PU-Tx; 2) the feedback from a cooperative sensing node located in the vicinity of the PU-Rx.

coordination among distributed CR transmitters. This distributed power-control algorithm can approximate the optimal solution without PU cooperation, central controller/monitor, or inter-SU message passing. However, the proposed algorithm is only effective for the *sum-utility* maximization, where the utility of each CR user only requires the CSI of itself. The utility function in [56] does not take other CR users' interference into account, which avoids the interference CSI estimation among CR users.

As we know, the formulated problem and the proposed algorithm in [56] have the drawback of inaccurate throughput characterization, which provides a *suboptimal* solution for CR networks. Moreover, the *sum-utility* maximization is not suitable for competitive and autonomous CR users. Our learning-based power control scheme differs from previous works in at least one of the following aspects: 1) this work provides a first solution for the *incomplete-information* power control for *competitive* and *autonomous* CR users; 2) different from [56], our work considers the interference among multiplier CR users (attacks the research issue b)) that provides an accurate throughput characterization for each CR user; 3) this work discloses the relationship between the incomplete-information power-control game and the classic complete-information power-control game; 4) the convergence and learning rate upper-bound is provided theoretically; 5) the algorithm is robust and adaptive to dynamic changing wireless environment and varying PUs' activities.

Learning-based power-control game has also been studied in other wireless networks. In [57], a non-cooperative power-control algorithm with repeated games was studied for ad hoc networks. The authors provided the important insight that a felicitous intelligent learning behavior with self-incentive dynamics could eventually converge to steady state with a satisfactory system performance. In [22], a distributed discrete power-control problem was formulated as an N-person nonzero sum game. The proposed stochastic learning power-control algorithm were proved to converge to a stable Nash equilibrium.

In our earlier work [58], we presented a framework for multiple CR users to perform distributed power control through reinforcement learning. In this research, we expand that

work and provide more descriptions of the learning mechanism for wireless environments, the properties and advantage of applying the Bush-Mosteller reinforcement scheme, and the convergence and learning rate analyses for the proposed power-control algorithm.

1.3 Literature Review: Capacity Analysis in Cognitive Radio Networks

Stimulated by the seminal work of Gupta and Kumar [8], the fundamental throughput capacity scaling for large-scale wireless ad hoc networks has become an active research topic in recent years, e.g., [8] [9] [10] [11] [12]. Compared to classical wireless ad hoc networks, the key distinguishing factors of the dynamic network topology, and the time and the location varying spectrum availability make the analysis of scaling laws for CRAHNs very challenging. The research on scaling laws, i.e., the asymptotic throughput capacity and delay, etc., for CRAHNs is still in the infancy phase [59], [35] [36] [37]. In [59], the scaling law is studied for single-hop CRNs. In [35] [36], the throughput capacity scaling is considered for the CRAHNs under the spectrum-sharing scheme of the *spectrum underlay* [2] that makes use of the *spatial domain*. This scheme enables SUs to coexist with PUs in the licensed spectrum bands while causing no excessive interference to PUs. It is shown that both the CRAHN and the primary network can simultaneously achieve the same throughput scaling as a stand-alone network: specifically, the per-node throughput capacity of the primary network and the CRAHN are $\Theta(W/\sqrt{m\log m})$[5] *bits/sec* and $\Theta(W/\sqrt{n\log n})$ *bits/sec* respectively. In [37], the throughput capacity of multi-channel wireless ad hoc networks is investigated in the presence of constraints on channel switching. Indicated in [37], the network scenario is applicable to CRAHNs where the SUs utilize a portion of the spectrum that is not being used by PUs (*multi-spectrum feature*). They introduce a channel-assignment model wherein a node can only switch between a set of f contiguous channels ($2 \le f \le c$), where c is the number of total subchannels. The per-flow capacity for

[5]Notation W stands for the available network bandwidths; n denotes the density of CRAHNs; m denotes the density of primary networks.

the adjacent (c, f) channel-assignment model is shown to be $\Theta(W\sqrt{\frac{f}{cn\log n}})$, where $f/c \leq 1$. As shown in [2], this spectrum-sharing scheme is referred as the *spectrum overlay* [5] that makes use of the *frequency domain*. Moreover, besides the spectrum-access schemes that make use of the spatial domain and the frequency domain, the spectrum overlay can also make use of the *time domain*, where SUs use the time slots based on the on-and-off activity of PUs' traffic loads [2](*time-varying feature*).

To the best of our knowledge, the scaling-law research which considers the important *time-varying feature* of the spectrum-sharing schemes in CRAHNs under the impact of *PU activity* in the spectrum overlay is still not touched in the research community.

1.4 Our Approaches and Thesis Outline

In Chapter 2, the decentralized behavior of power and rate control in wideband CRNs is considered, where each SU competes for radio resources in the wideband spectrum-underlay CRNs. The problem of optimal power- and rate-control is investigated for each individual SU by using the non-cooperative game theory. The formulated problem satisfies the following two operational constraints in CRNs : (1) the total amount of interference power caused by SUs to PUs must not exceed a predefined threshold; (2) for each SU, the received SINR must exceed a predefined threshold to guarantee its QoS. The utility function considered in this research measures the throughput (transmission rate) per transmit power. It guarantees energy efficiency and fairness among SUs. The utility function is especially suitable for the feature of energy-constrained, low-power and competition-based data communication of SUs.

The formulated problem is a *constrained game-theoretic* problem [45]. By introducing pricing into the utility functions, the interference power caused from each SU to PUs is charged by price, which is a more intuitive representation compared to the original problem. Meanwhile, as we know, pricing is also an effective tool for radio resource management because of its ability to guide user's behavior toward a more efficient operating point [42].

That means, on one hand, that we can use the pricing to study the interference power caused by SUs; on the other hand, the net utilities or the satisfactions experienced by SUs will be greatly improved when a proper pricing mechanism is introduced. Based on these analyses, the non-cooperative joint-power and rate-control game with interference-power pricing is proposed. A detailed analyses of the existence, uniqueness, and Pareto efficiency of the Nash equilibrium for the proposed game are presented. Then, a joint-power and rate-control algorithm is proposed to find the optimal transmission rates and transmit powers for SUs. Moreover, the process of seeking the best interference-power pricing factors for SUs that leads to the optimal performance is discussed. Then, an algorithm is designed to find the best pricing factors. In the numeral results sections, the performance of the proposed joint-power and rate-control algorithm is investigated.

In Chapter 3, the multi-user power-control problem is studied for CRNs overall multiple channels in the dynamic fading environments through a game-theoretical approach. This research provides the fundamental performance traits of multiple CR users power control in the present of multiple PUs in fading environments. in fading environments. Unlike in [16] and [48] where information-theoretical capacity is considered as the objective function, a more practical and general performance merit in [30] and [26] is considered in this research: the target SINR of each individual users. It can be used to achieve both a specific transmission rate (i.e. by using different coding and modulation schemes) in practical and the information-theoretical capacity (like [16] and [48]) in theory. Due to channel fading, the target SINR of each CR user can not always be achieved over time as in [30] and [26]. There exists channel outage probability related to both CR and PU transmissions, which is given by the probability that the SINR falls below a predetermined threshold. As we have noticed, the formulated problem can be effectively transformed into a convex-optimization problem by applying the geometric programming [60]. A partial dual-decomposition approach is applied with layered structure for the constrained game. The properties of the Nash Equilibrium for the proposed game are thoroughly investigated, and an asynchronous

distributed power-control algorithm that converges to the Nash Equilibrium is proposed for the robust implementation

In Chapter 4, for the learning-based power-control problem, it is formulated as an incomplete-information non-cooperative game in CR networks, where the interference channel and power strategy information among CR users and the interference channel information from CR users to PUs are not available. The power-control game as an interaction process is modeled by repeat games [61]. The proposed new framework requires CR users to be "cognitive" enough as in [56] that could decode the link control information from PUs' feedback channels. During every stage of the repeated game, each CR user will only utilized the interference information explored from PUs' control link and the transmission rate (by trying certain power control strategy) obtained in the last stage to decide the power control strategy in this stage, but there is no other information. A reinforcement-learning approach is designed for the repeated power-control game. The proposed algorithm is suitable for a wide scope of practical CR networks. To name a few, the networks scenarios are: 1) there are no creditable CR users; 2) no infrastructure-supported CR networks; 3) CR users are not willing to exchange their private information; 4) the large network deployment condition, where channel information exchange does not scales well; 5) fierce channel varying environments, where channel estimation among CR users might not be easily updated; 6) PUs' activity statistics are unknown, e.g. PU's on-and-off distribution are unknown, the interference-power constraints of PU might change during the time (e.g. for mobile PUs).

Similar to the classic game-theoretical approach, the objective of the proposed power-control game is as follows: on one hand, CR users need to compete for maximizing their own average utility in the repeated game; on the other hand, each individual CR user needs to satisfy the interference power constraints imposed by PUs during the course of the game. The formulated power-control problem is a *constrained* repeated game with learning automaton. The solution of this repeated game is to choose the optimal transmit power

strategy at each stage that maximizes the average utility and conform the interference power constraints from PUs in a long run. As a learning automaton, the CR user will adjust the mixed transmit power strategies according to its own realization of utility and the explored interference level information at every time step, but no other information. The power-control game is proved to be asymptotically equivalent to the classic game with complete information. The properties of existence, diagonal concavity and uniqueness for the matrix game from the view point of each individual CR user are studied. One of the best known stochastic models of reinforcement learning, the Bush-Mosteller reinforcement learning procedure [62], is used to designed the power-control algorithm, and the properties of convergence and learning rate of the algorithm are analyzed. Finally, the learning-based power-control algorithm is implemented for simulations.

In Chapter 5, the fundamental scaling-law problem that how the capacity and delay of CRAHNs scale under the impact of PU activity is analyzed. To make this research of practical value, the primary network is assumed to be a collection of independently placed transceiver pairs like the cellular or broadcasting networks, where each transceiver pair is the licensed spectrum holders. The CRAHN is assumed to be a static ad hoc network as in [12][13][14] that distributed in the region of the primary network. The performance of CRAHN is suffered by PU activity, and there is a large space of the scheduling policies or protocols design for the CRAHNs that leads to the optimal network performance. The aim of this research is to study under what kind of typical scheduling policy or protocol the optimal scaling performance of CRAHNs (delay-optimal or throughput-optimal) can be achieved, and how the primary network deployment impact them.

We assume that there are m primary transmit-receive pairs with on-and-off probabilities P_{on} and P_{off} independently distributed in the region of a CRAHN with n secondary nodes[6] and the total bandwidth of W bits per second. Our work differs from the previous studies in the following critical points.

[6]the terms ”nodes” and ”users” are used interchangeably in this work.

1) It is observed in [35] [36] [37] that the scaling results for CRAHNs have no obvious relation with the primary network deployment condition when secondary nodes take advantage of spectrum opportunity in either the frequency or the spatial domain. However, by utilizing the time domain spectrum opportunity granted by PU activity, in this work the scaling performance of CRAHNs are tightly related to the primary network topology. Specifically, the primary nodes' density m, the transmission range and the on-and-off probabilities P_{off} (or P_{on}) are exactly embodied in the scaling results of CRAHNs. These results are of strong theoretical meanings. They bridge the relation between primary networks and CRAHNs, which provides an initial and fundamental understanding for the two co-existence networks. These results can be served as design guidance of CRAHNs in the present of PU activity.

2) By considering the PU activity, it is found that the behavior of secondary traffic in CRAHNs becomes heterogeneous in order to achieve the maximum throughput capacity. Specifically, this means in classical ad hoc wireless networks [8] [10] and CRAHNs [35] [36], there is always traffic load for each source-destination pairs in any specified time and location. However, when PU activity is involved, we show that the secondary nodes need to buffer their traffic in order to achieve the maximum throughput capacity. As a result, there may not always be traffic in CRAHNs at a specific time and location. Based on this reason, the queuing and transmission delays must be involved into end-to-end delay computation. In [10] and [36], only the delay arising out of multi-hop operation is considered for classic wireless ad hoc networks and CRAHNs, which is proportional to the number of hops between the source and destination nodes.

To study the impact of primary networks' topology on scaling laws of CRAHNs, we first consider a typical primary network model: *regular dense primary network*[7], where primary transmitters are regularly placed with equal transmission range (primary receivers are therein) in the network region. In this case, the primary network is very dense and the

[7]Detailed definitions and illustrations are available in Sections III,IV, V.

primary nodes' transmission ranges are contingent with each other with high probability, where the secondary nodes of CRAHNs are nearly all covered within the primary nodes' transmission ranges. This scenario achieves the worst performance in CRAHNs. Further, we consider another typical primary network model: *regular sparse primary network*[3], where primary nodes' density is sparse and the transmission range is small compared with the regular dense primary network. In this case, there are always secondary nodes that are uncovered by primary nodes' transmission ranges and finding routing paths outside of these uncovered regions for each secondary node is proved to be feasible. Finally, we generalize our results to the *random dense* and *sparse primary networks*[3] where the location of primary nodes can be randomly distributed.

In Chapter 6, conclusion is drawn for the thesis, where the main results of power control and capacity scaling of CRNs are presented.

CHAPTER 2

JOINT-POWER AND RATE-CONTROL GAME VIA PRICING IN WIDEBAND COGNITIVE RADIO NETWORKS

2.1 System Model

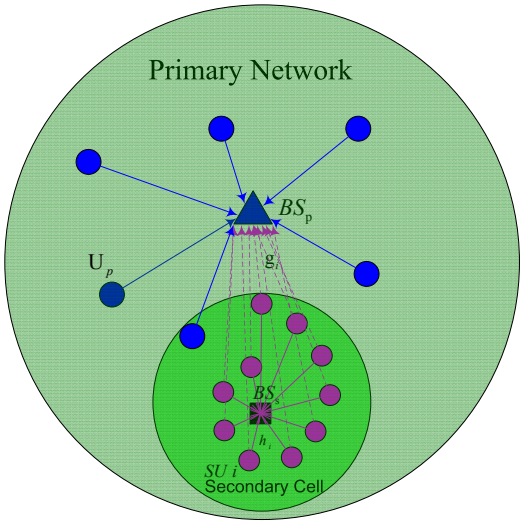

Figure 2.1. System model of the primary and cognitive radio networks.

The system model is a heterogeneous network shown in Fig.2.1. It is a large 3G wireless cellular primary network that contains a secondary cell. This is a representative scenario of SUs communications in the spectrum-underlay scheme. The following typical communication scenario is considered in our discussion: a group of secondary mobile stations or SUs randomly distributed in the cell are transmitting data to the secondary base station; PUs are now in the uplink communication, where primary base station measures the interfering power from SUs. For the other possible scenarios, such as secondary downlink transmission and/or primary downlink transmission, the principal of secondary transmission making interference to primary transmission is the same.

Assume that the PUs, denoted as U_p, are communicating with the primary base station (BS), denoted as BS_p. They transmit on a licensed spectrum W. The CRN is a single-cell spread-spectrum wireless network without the license of the spectrum which

lies in the communication range of the 3G primary cell (network). There are N SUs that use the spread-spectrum scheme transmitting data to the secondary BS, denoted as BS_s. We assume all SUs use a spread-spectrum signaling format to share the available spread-spectrum bandwidth (unit is HZ). In this model, the available spread-spectrum bandwidth is W, which is exactly the same usable spectrum bands for the licensed PUs. For the uplink scenario, all the SUs must restrain their transmit powers in order not to cause excessive interference to the PUs' BS BS_p.

For secondary data communication, there are N active CR links, each of which interferes with the other $N-1$ SUs. The SUs are mobile terminals in the cell, and the channel gain between SU i and BS_s is denoted by $h_i\,(i = 1, 2, ..., N)$. The secondary base station BS_s help sending pilot or training singling to estimate the channel gains between SUs and itself. The SINR of a given SU i as referred in [42] [44], is given by

$$\gamma_{\mathrm{i}}\,(r_i, p_i) = \frac{W}{r_i} \cdot \frac{h_i p_i}{\sum_{j=1, j\neq i}^{N} h_j p_j + \sigma^2},\ \textit{for all}\ \mathrm{i} = 1, 2, ..., N. \tag{2.1}$$

The parameters p_i and r_i are the respective transmit power and transmission rate of the SU i. These formulas present that all SUs use a spread-spectrum signaling format over the whole shared band W.

For the impact of SUs' transmissions on PUs, primary base station BS_p can help estimating the channel gains from SUs to itself, measuring the instantaneous total interference power from all SUs, and dispatching the interference-channel gains and interference-power threshold to the SUs. The SUs will control their transmit powers based on these information. Let $g_i\,(i = 1, 2, ..., N)$ be the channel gain of the ith SU's transmitter to the PUs' BS BS_p. Define p_i as the transmit power of SU i and $\mathbf{p} = (p_1, p_2, ..., p_N)$ as the transmit-power vector of the N SUs. The interference power received by the PUs' BS BS_p is $\sum_{i=1}^{N} \mathrm{g}_i p_i$. To facilitate the secondary cellular network coexistence with PUs, the interference power from all SUs should be lower than certain tolerable threshold. In this paper, we use $\mathrm{P_{th}}$ to denote the interference-power threshold. So the *interference-power constraint* of SUs in the CRN

is given by

$$\sum_{i=1}^{N} g_i p_i \leq \mathrm{P_{th}}. \tag{2.2}$$

2.2 Problem Formulation

In the heterogeneous network, as depicted in Fig.1, the behavior of N SUs can be perfectly modeled as a game: SUs in the secondary cell will compete selfishly to make best use of network resource (e.g. spectrum/rate and power) provided by PUs. On one hand, each SU will try its best to achieve a high SINR to guarantee a reliable data transmission; on the other hand, under certain target SINR requirement, each SU will use a low transmit power and achieve a high transmission rate to maximize its own utility. In this paper, since we focus on distributed schemes, we will concentrate on non-cooperative games and assume each SU behaves rationally to maximize its own utility and avoid inference to PUs.

Let $\mathbb{G} = [\mathbb{N}, \{\mathbb{P}, \mathbb{R}\}, \mathbb{U}(\cdot)\}]$ denote *the non-cooperative power- and rate-control game* (NPRG) in the CRN, where $\mathbb{N} = \{1, 2, ..., N\}$ is the set of players. The players in this game correspond to the N SUs. $\mathbf{P}_i$ is the transmit-power strategy space of SU i. $\mathbf{R}_i$ is the transmission-rate strategy space of SUs. Each SU selects a rate $r_i \in \mathbf{R}_i$ and a power $p_i \in \mathbf{P}_i$. Here we assume that the strategy space $\mathbf{R}_i$ and $\mathbf{P}_i, i = 1, 2, ..., N$ for each SU are compact and convex sets with maximum and minimum constraints. For SU i, we consider rate strategy space $\mathbf{R}_i$= [$\mathrm{r_i^{min}}$,$\mathrm{r_i^{max}}$] and power strategy space $\mathbf{P}_i$= [$\mathrm{p_i^{min}}$,$\mathrm{p_i^{max}}$], where $\mathrm{r_i^{min}}$, $\mathrm{r_i^{max}}$, $\mathrm{p_i^{min}}$ and $\mathrm{p_i^{max}}$ are the minimal and maximal transmit power and transmission rate imposed on each SU. These values correspond to dynamic ranges labeled in the SU transmitters, e.g., the maximal transmission rate is related to the maximal order of coding and modulation schemes the SU can support, the maximal transmit power is the peak power the SU transmitter can support. Let the rate vector of the SUs system denotes as $\mathbf{r} = (r_1, r_2, ..., r_N)$ and the power vector as already defined, $\mathbf{p} = (p_1, p_2, ..., p_N)$. The outcome of the game in terms of selected powers and rates of all SUs can be given as $\mathbb{P} = \mathbf{P}_1 \times \mathbf{P}_2 \times ... \times \mathbf{P}_\mathrm{N} = \left\{\mathbf{p} | \mathrm{p_i^{min}} \leq p_i \leq \mathrm{p_i^{max}}\right\}$ and $\mathbb{R} = \mathbf{R}_1 \times \mathbf{R}_2 \times ... \times \mathbf{R}_\mathrm{N} = \left\{\mathbf{r} | \mathrm{r_i^{min}} \leq r_i \leq \mathrm{r_i^{max}}\right\}$.

The utility function of the SU i is denoted as $u_i(\cdot)$, where $u_i(\cdot) \in \mathbb{U}(\cdot)$ is the utility-function space.

In the non-cooperative power- and rate-control game, each SU seeks to choose its power p_i and rate r_i in such a way as to maximize its utility u_i. So the NPRG can be formally expressed as

$$\mathbf{NPRG} \max_{p_i \in \mathbf{P}_i, r_i \in \mathbf{R}_i} u_i(\mathbf{p}, \mathbf{r}), \forall i \in N. \tag{2.3}$$

2.2.1 Utility Function

The selection of a proper utility function has a great impact on the nature of the NPRG game in the CRN. For data-oriented, energy-constrained and low-power SUs in the 3G wireless networks, a good choice for the utility function is the one that measures the number of bits successfully transmitted per joule of consumed energy. Note that, on one hand, a higher SINR at the receiver side will result in a lower bit-error-rate and thus increase SU's utility. However, on the other hand, to achieve a high SINR level often requires SU to transmit at a high power which will cause interference to others SUs and PUs. This tradeoff can be described by defining the utility function of a SU as the ratio of its throughput to its transmit power, i.e,

$$u_i = \frac{T_i}{\mathrm{p}_i}. \tag{2.4}$$

Throughput here is the net number of bits that are transmitted without error per unit time. It can be expressed as follows:

$$T_i = r_i f(\gamma_i), \tag{2.5}$$

where r_i and γ_i are the transmission rate and SINR for the ith SU. $f(\gamma_i)$ is the *efficiency function* which represents the *frame success rate* (FSR) [63]. The meaning of efficiency function is that if a frame has one or more bit errors, it will be retransmitted. The utility function in (4) was introduced in [63] [64]. Its unit is *bits/joule*. It is a widely used model to characterize the tradeoff between power interference and networks throughput. It has been used by others in different scenarios (for example, [41], [42], [65]). Note that $f(\gamma_i)$ depends

on the specific data-transmission purposes, such as the modulation scheme, coding, frame size and fairness issue, etc. In this paper, with the purpose of guaranteeing fairness among these competitive SUs, we adopt the logarithmic efficiency function as follows to study the joint-power and rate-control game. As indicated in [66], logarithmic functions can guarantee proportional-fairness among all the terminals users.

$$f(\gamma_i) = \ln(K\gamma_i). \tag{2.6}$$

In (6), parameter K can be adjusted to qualify the QoS requirement (we will show this point in the following discussion immediately). Combining (4), (5) and (6), the utility function of the ith SU is given by

$$\mathrm{u}_i = r_i \frac{\ln(K\gamma_i)}{p_i}. \tag{2.7}$$

For rate allocation in the problem (3), the SU i need to find a transmission rate r_i^* from the strategy space $\mathbf{R}_i$ to maximize its utility function u_i. The maximum rate achieves at the point where the partial derivative of u_i with respect to i is zero, $\partial \mathrm{u}_i/\partial r_i = 0$. So the condition for maximizing transmission rate is

$$\frac{\partial \mathrm{u}_i}{\partial r_i} = \frac{1}{p_i}(\ln(K\gamma_i) - 1) = 0. \tag{2.8}$$

For power allocation, the SU i needs to find a transmission power p_i^* from the strategy space $\mathbf{P}_i$ to maximize its utility function u_i ($\partial \mathrm{u}_i/\partial p_i = 0$). Through calculation we find it's the same condition as (8).

The condition for optimal power- and rate-control can be used to obtain the parameter K. According to (7), K can be obtained as:

$$K = \frac{e}{\gamma_i^{tar}}, \tag{2.9}$$

where γ_i^{tar} is the target SINR for the SU i as its QoS requirement. The value of K in (9) is calculated when there is no bit error occurs in the transmission. However, in a real network, this will not be guaranteed. A more practical way of determining K is to set the efficiency

function $f(\gamma_i)$ equals to the correct reception probability P_c.

$$f(\gamma_i) = \ln\left(K\gamma_i^{tar}\right) = P_c. \tag{2.10}$$

Solve this equation, the parameter K is obtained as follows:

$$K = \frac{e^{P_c}}{\gamma_i^{tar}}. \tag{2.11}$$

The advantage of parameter K is that the target SINR can be determined by adjusting its value. Numerical simulation shows that by changing the value of K, SUs can obtain a different equilibrium SINR. Thus, with predefined value of K, SUs can play the joint-power and rate-control game to achieve their target QoS. This means the SINR of each SU must be equal to or greater than γ_i^{tar} to guarantee a reliable transmission,

$$\gamma_i \geq \gamma_i^{tar}. \tag{2.12}$$

Based on the above discussion, the joint-power and rate-control problem for the SUs system can be formulated as follows:

$$\begin{aligned} \max_{p_i \in P_i, r_i \in R_i} \quad & u_i(p, r) = \tfrac{r_i \ln(K\gamma_i)}{p_i}, \forall i \in N \\ \textit{subject to} \quad & \mathrm{p_i^{min}} \leq p_i \leq \mathrm{p_i^{max}}, \forall i \in N \\ & \mathrm{r_i^{min}} \leq r_i \leq \mathrm{r_i^{max}}, \forall i \in N \\ & \sum_{i=1}^{N} g_i p_i \leq \mathrm{P_{th}} \\ & \gamma_i \geq \gamma_i^{tar}. \end{aligned} \tag{2.13}$$

The problem (13) has multiple objective functions with non-linear coupled constraints, which is a constrained game-theoretical problem [67]. It is not easy to approach through standard optimization theories. Motivated by the concept *pricing* from the game theory [45], by introducing it into the utility functions, the net utility for each SU can be calculated. Note that an efficient pricing mechanism will make a distributive decisions among SUs that are compatible with the overall system efficiency by not only encouraging efficient resources share provided by PUs, but also avoiding interference power to PUs.

Since the interference power from SUs to PUs is a linear combination of the power vector $\mathbf{p} = (p_1, p_2, ..., p_N)$ and the channel-gain vector $\mathbf{g} = (g_1, g_2, ..., g_N)$ in (2), we can use the linear pricing function for this game.

Now, a non-cooperative game with pricing is defined, in which the pricing is used to characterize the negative effect from each SU to PUs. Let $\mathbb{G}_{\mathrm{p}} = [\mathbb{N}, \{\mathbb{P}, \mathbb{R}\}, \{\mathbb{U}^{\mathrm{p}}(\cdot)\}]$ represent *the non-cooperative joint-power and rate-control game with interference-power pricing* (NPRGP). $u_i^p(\cdot) \in \mathbb{U}^{\mathrm{p}}(\cdot)$ is the utility function for the NPRGP. The utility function is given by

$$u_i^p(\mathbf{p}, \mathbf{r}) = u_i(\mathbf{p}, \mathbf{r}) - c_i(p_i), \tag{2.14}$$

where $c_i(p_i)$ is the pricing function for SU $i \in N$. The linear pricing model of $c_i(\mathrm{p})$ is given by

$$c_i(p_i) = c_i \cdot p_i. \tag{2.15}$$

The pricing factor c_i ($c_i \geq 0$) should be adjusted such that the self interest of each SU leads to the overall improvement of the system. Meanwhile, by adjusting c_i for each SU, the overall improved system utility must conform the interference-power constraint in (2). So, the utility function of each SU in the NPRGP can be expressed as

$$\mathbf{NPRGP} \max_{p_i \in \mathbf{P}_i, r_i \in \mathbf{R}_i} u_i^p(\mathbf{p}, \mathbf{r}) = \frac{r_i \ln(K\gamma_i)}{p_i} - c_i \cdot p_i, \forall i \in N. \tag{2.16}$$

Moreover, in the following discussion, it will be shown that regardless of how to choose the pricing factor c_i and whether it meets the interference-power constraint or not, the Nash equilibrium for the NPRGP always exist.

2.3 Nash equilibrium and Pareto Efficiency

In this section, a power- and rate-control algorithm is derived for the NPRGP game in which each SU aims to maximizing its net utility. This can be achieved by checking the first-order condition of the optimal point. For rate allocation, the optimization problem for the SU i is to find the rate level r_i^* from the strategy space $\mathbf{R}_i$ to maximize the utility function. For

power allocation, the optimization problem for the SU i is to find the power level p_i^* from the strategy space $\mathbf{P}_i$ to maximize the utility function.

2.3.1 Nash Equilibrium of the NPRGP

Definition 1: Let $\mathbf{A}_i$ be a strategy space, where $a_i \in \mathbf{A}_i$. A vector $\mathbf{a}^* = \left(a_1^*, a_2^*, ..., a_N^*\right)$ in the strategy spaces is a Nash equilibrium, such that no user can improve its own utility unilaterally, that is,

$$\mathrm{u}_i\left(a_i^*, \mathbf{a}_{-i}^*\right) \geq \mathrm{u}_i\left(a_i, \mathbf{a}_{-i}^*\right), \forall a_i \in \mathbf{A}_i, and, \forall i \in N, \tag{2.17}$$

where $\mathbf{a}_{-i}^* = \left(a_1^*, ..., a_{i-1}^*, a_{i+1}^*, ..., a_N^*\right)$.

A Nash equilibrium is a stable outcome of a game. At Nash equilibrium point, no user has any incentive to change its strategy with its own action.

Theorem 1 (*Existence*): There exists a Nash equilibrium of the transmission rates $\mathbf{r}^* = \left(r_1^*, r_2^*, ..., r_N^*\right)$ in the NPRGP game $\mathbb{G}_\mathrm{p} = [\mathbb{N}, \{\mathbb{P}, \mathbb{R}\}, \{\mathbb{U}^\mathrm{p}(\cdot)\}]$, $\forall i \in N$ for the CRN.

Proof The conditions for the existence of a Nash Equilibrium are given in [42]:

(1) $\mathbf{R}_i$ is a nonempty, convex, and compact subset of some Euclidean space $\mathbb{R}^\mathrm{N}$;

(2) $u_i(\mathbf{r})$ is continuous in $\mathbf{r}$, and quasi-concave in r_i.

Because each SU is assumed to have a strategy space that is an interval between the maximum and minimum rates. So the first condition on the strategy space $\mathbf{P}_i$ is satisfied.

The first-order partial derivative of $u_i^p(p, r)$ with respect to r_i is given by

$$\frac{\partial u_i^p}{\partial r_i} = \frac{1}{p_\mathrm{i}}\left(\ln\left(K\gamma_i\right) - 1\right). \tag{2.18}$$

Relaxing the upper bound of the rate region to find the possible r_i^*, the best response value r_i^* is obtained when the following condition is satisfied:

$$\left.\frac{\partial u_i^p\left(r_i, \mathbf{r}_{-i}\right)}{\partial r_i}\right|_{r_i = r_i^*} = 0. \tag{2.19}$$

To show a function is quasi-concave, it is sufficient to show it is concave, and we use the second derivative test for this.

$$\frac{\partial^2 u_i^p}{\partial r_i^2} = \frac{1}{p_i}\frac{\partial \ln\left(K\gamma_i\right)}{\partial r_i} = -\frac{1}{p_i r_i}. \tag{2.20}$$

In (20), since $\partial^2 u_i^p \big/ \partial r_i^2 < 0, \forall i \in N$, u_i is a concave function of r_i. Therefore, it guarantees the existence of a Nash equilibrium.

In the next, the uniqueness of the Nash equilibrium of the game $\mathbb{G}_{\mathrm{p}}$ is proved.

Proposition 1 : For the NPRGP game $\mathbb{G}_{\mathrm{p}}$, the best response function of the *i*th user given the transmission-rate vector of other users $\mathbf{r}_{-i}$, is given by: $\rho_i(\mathbf{r}_i) = \min\left(r_i^*, \mathrm{r}_{\mathrm{i}}^{\max}\right), \forall i \in N$, where $\mathrm{r}_{\mathrm{i}}^{\max}$ is the maximum transmission rate in the *i*th user's strategy space $\mathbf{R}_i$.

Proof the best response function $\rho_i(\mathbf{r}_i)$ of the *i*th user is defined as the best action that user *i* can take to attain the maximum utility given the other users' strategy $\mathbf{r}_{-i}$. Formally, user *i*'s best response $\rho_{\mathrm{i}} : \mathbf{R}_{-i} \rightarrow \mathbf{R}_i$ is the mapping that assigns to each $\mathbf{r}_{-i} \in \mathbf{R}_{-j}$ the set.

$$\rho_i(\mathbf{r}_{-i}) = \left\{ r_j \in \mathbf{R}_j : u_i(r_i, r_{-i}) \geq u_i(r_i', r_{-i}), \forall r_i' \in \mathbf{R}_i \right\}, \tag{2.21}$$

where this set contains only one point [45]. Therefore, r_i^* is the unconstrained maximizer of the target utility function u_i.

$$r_i^* = \arg\max_{r_i \in R^+} u_i^p. \tag{2.22}$$

Moreover, from (20), we have $\partial^2 u_i^p \big/ \partial r_i^2 < 0, \forall r_i \in R^+$. It implies that the maximum is unique. Note that when r_i^* is not a feasible solution in the strategy space $\mathbf{R}_i$, that is $r_i^* \notin \mathbf{R}_i$, user *i* will transmit at the maximum rate $\mathrm{r}_i^{\max}$ since the target function is increasing on the set $\left\{ r_i : r_i < r_i^* \right\}$. This implies that $r_i = r_i^*$ is the best response of user *i*.

The following theorem, proved in [41], guarantees the uniqueness of the Nash equilibrium of game $\mathbb{G}_{\mathrm{p}}$.

Theorem 2: If a power-control algorithm with a *standard* best response function has a Nash equilibrium, then this Nash equilibrium is unique.

A function is said to be *standard*, if it satisfies the following three properties:

(1) **positivity**: $f(\mathbf{x}) > 0$

(2) **monotonicity**: if $\mathbf{x} \geq \mathbf{x}'$, then $f(\mathrm{x}) \geq f\left(\mathbf{x}'\right)$

(3) **scalability**: for all $\mu > 1, \mu f(\boldsymbol{x}) > f(\mu \mathbf{x})$

where $\mathbf{x}$ is a Nash equilibrium and $f(\mathbf{x})$ is the best response function. According to theorem 2, we can prove the following lemma for the transmission-rate control similarly.

Lemma 1 *(Uniqueness)*: In the NPRGP game $\mathbb{G}_{\mathrm{p}}$, the best response transmission-rate vector, given by

$$\rho(\mathbf{r}) = (\rho_1(\mathbf{r}), \rho_2(\mathbf{r}), ..., \rho_N(\mathbf{r})),$$

is a standard function. Therefore, by theorem 2, there has a unique Nash equilibrium $\mathbf{r}^* = \left(r_1^*, r_2^*, ..., r_N^*\right)$.

Proof from theorem 1, we know that there exists a Nash equilibrium $\mathbf{r}^*$. By definition, the Nash equilibrium has to satisfy $\mathbf{r}^* = \rho(\mathbf{r})$, where $\rho(\mathbf{r}) = (\rho_1(\mathrm{r}), \rho_2(\mathbf{r}), ..., \rho_N(\mathbf{r}))$ is the best response vector of all users. According to theorem 2, the uniqueness can be proved if we show that the best response function $\rho(\mathbf{r})$ is standard. The three properties of standard function can be easily verified for $\rho(\mathbf{r})$. It is shown in [63] that the fixed point $\mathbf{r}^* = \rho(\mathbf{r})$ is unique for a standard function. Therefore, the Nash equilibrium is unique.

Theorem 3 (Existence and Uniqueness): If c_i in the pricing function satisfies $c_i = \tilde{c}_i$, where $\tilde{c}_i = r_i\left[\frac{1-\ln(K\gamma_i)}{p_i^2}\right]$, there exists a unique Nash equilibrium $\mathbf{p}^* = \left(p_1^*, p_2^*, ..., p_N^*\right)$ of the transmit power in the NPRGP game $\mathbb{G}_{\mathrm{p}} = [\mathbb{N}, \{\mathbb{P}, \mathbb{R}\}, \{\mathbb{U}^{\mathrm{p}}(\cdot)\}]$ for the CRN.

Proof similar to theorem 1, for the proof of the existence of the Nash Equilibrium, the first condition of theorem 1 about the convexity and compactness of the strategy space $\mathbf{P}_i$ is satisfied obviously, since $\mathbf{P}_i = [\mathrm{p}_i^{\min}, \mathrm{p}_i^{\max}]$ is a compact interval.

The first-order partial derivative of $u_i^p(p, r)$ with respect to p_i is given by

$$\frac{\partial u_i^p}{\partial p_i} = r_i\left[\frac{1-\ln(K\gamma_i)}{p_i^2}\right] - c_i. \tag{2.23}$$

Relaxing the upper bound of the rate region to find the possible p_i^*, the best response value p_i^* is obtained when the following condition is satisfied:

$$\left.\frac{\partial u_i^p(p_i, \mathbf{p}_{-i})}{\partial p_i}\right|_{p_i = p_i^*} = 0. \tag{2.24}$$

To show a function is quasi-concave, we have

$$\frac{\partial^2 u_i^p}{\partial p_i^2}=r_i\frac{\ln(K\gamma_i)-1}{p_i^3}. \tag{2.25}$$

We know that the probability of correct reception is always less than or equal to 1 (As indicated in (10), $P_c \leq 1 \Rightarrow \ln(K\gamma_i) \leq 1$.) Substitute this into the above (25), we have $\partial^2 u_i^p \Big/ \partial p_i^2 \leq 0$, which implies that u_i is a quasi-concave function of p_i optimized on the convex set $\mathbf{P}_i$. This proves condition 2, which guarantees the existence of a Nash equilibrium.

Now we prove the uniqueness of the Nash equilibrium of transmit powers in the NPRGP game $\mathbb{G}_{\mathrm{p}}$. Set (23) to be zero, we get

$$\tilde{c}_i = r_i\left[\frac{1-\ln(K\gamma_i)}{p_i^2}\right], \tag{2.26}$$

where $(1-\ln(K\gamma_i))\Big/p_i^2$ is a strictly decrease function with respect to p_i. If $c_i = \tilde{c}_i$, it satisfies that

$$\left.\frac{\partial u_i^p(p_i,\mathbf{p}_{-i})}{\partial p_{\mathrm{i}}}\right|_{p_i=p_i^{min}} > 0. \tag{2.27}$$

And,

$$\left.\frac{\partial u_i^p(p_i,\mathbf{p}_{-i})}{\partial p_i}\right|_{p_i=p_i^{\max}} < 0. \tag{2.28}$$

So there exists one and only one $p_i^* \in [p_i^{\min}, p_i^{\max}]$, such that

$$\frac{\partial u_i^p(p_i,\mathbf{p}_{-i})}{\partial p_i}\begin{cases} >0 & p_i^{\min} < p_i < p_i^* \\ =0 & p_i = p_i^* \\ <0 & p_i^* < p_i < p_i^{\max} \end{cases}. \tag{2.29}$$

Hence, there exists a unique best response strategy p_i^* for SU i, if $c_i = \tilde{c}_i$. In this case, a unique Nash equilibrium exists.

As in the theorem 1, considering the upper bounds of SUs' transmit powers, the best response strategy of the transmit power of the SU i is

$$p_i^*(\mathbf{p}_{-i}) = \min\left(p_i^*, p_i^{\max}\right), \forall i \in N, \tag{2.30}$$

where $\overset{*}{p}_i = \arg\max_{p_i \in R^+} u_i^p$. It still guarantees the existence of a unique Nash equilibrium.

2.3.2 Pareto Optimality of NPRGP

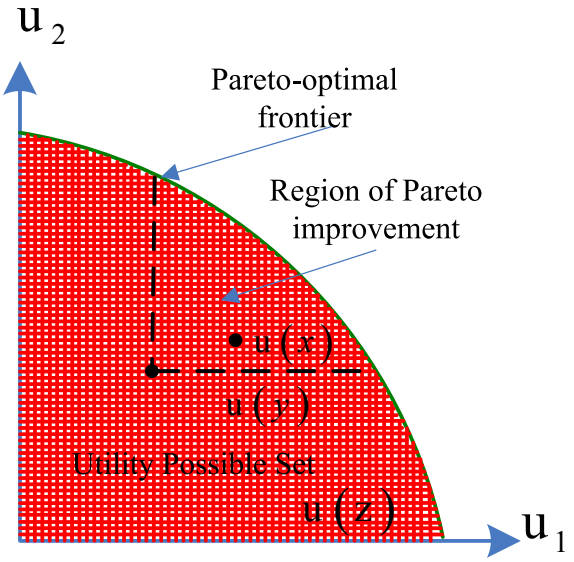

Figure 2.2. Power vector x Pareto dominates power vector y, and z is Pareto optimal.

The Nash equilibrium discussed in last subsection is the solution to the power- and rate-control problem where no SU can increase its utility any further unilaterally. Indeed, it is well known that in general the Nash equilibria are inefficient [45]. A resource-allocation policy is said to be more efficient (or called "*Pareto dominant*"), if it is possible to increase the utility of some of the terminals without hurting any other terminal. This efficiency is usually referred as *Pareto efficient* or *Pareto optimal*. The definition is as follows.

Definition 2: A vector $\hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2, ..., \hat{a}_N)$, whether transmit power or transmission rate, Pareto dominates another vector $\mathbf{a}$, if, for all $i \in N$, $u_i(\hat{\mathbf{a}}) \geqslant u_i(\mathbf{a})$ and for some $i \in N$, $u_i(\hat{\mathbf{a}}) > u_i(\mathbf{a})$. Furthermore, a vector $\tilde{\mathbf{a}} = (\tilde{a}_1, \tilde{a}_2, ..., \tilde{a}_N)$ is Pareto Optimal (efficient) if exists no other vector $\mathbf{a}$ in the strategy space such that $u_i(\mathbf{a}) \geqslant u_i(\tilde{\mathbf{a}})$ for all $i \in N$, and $u_i(\mathbf{a}) > u_i(\tilde{\mathbf{a}})$ for some $i \in N$.

Fig.2.2 explains the concept of Pareto dominance and Pareto optimality on a generic utility possibility set. In Fig.2.2, there are two terminals in the game, and their strategy sets are mapped to the utility possibility set shown in the shaded area. Take the power-control

game as an example, any power vector that provides a Pareto improvement with respect to $\mathbf{y}$ results in nondecreasing changes in individual utilities, $u_i(\mathbf{y})$. From the figure, we can observe that $\mathbf{x}$ is such a power vector. We can also call $\mathbf{x}$ as the *Pareto-preferred* power vector when compared to $\mathbf{y}$. The distinction of the concepts *Pareto dominance* and *Pareto optimality* is that: Pareto-optimal power allocations *do not* necessarily Pareto dominate *all* other power vectors. For example, compare the utilities obtained by $\mathbf{y}$ and $\mathbf{z}$ in Fig.2.2, $\mathbf{z}$ is a Pareto-optimal power allocation. However, since $u_2(\mathbf{z}) < u_2(\tilde{\mathbf{y}})$, $\mathbf{z}$ does not Pareto dominate $\mathbf{y}$, regardless of the fact that $u_1(\mathbf{z}) > u_1(\tilde{\mathbf{y}})$.

In this paper, the power and rate vectors that improve utilities, in the Pareto-sense, with respect to the Nash equilibrium are referred to as the NPRGP-*dominant*. The focus of this section and the next section on pricing is to seek the NPRGP-dominant power and rate allocations. The following theorem testifies the Pareto optimality (efficiency) of the equilibrium point $(\mathbf{r}^*, \mathbf{p}^*)$ in the NPRGP game.

Theorem 4: The Nash equilibrium point $(\mathbf{r}^*, \mathbf{p}^*)$ of the NPRGP game is Pareto optimal.

Proof mathematically speaking, for $\mathbf{r}$ in $\mathbb{G}_{\mathrm{p}}$, $\nexists \mathbf{r}^o = \left(r_1^o, r_2^o, ..., r_N^o\right) : u_i(\mathbf{r}^o) \geq u_i(\mathbf{r}^*), \forall i \in N$ and $u_j(\mathbf{r}^o) \geq u_j(\mathbf{r}^*)$, for some $j \in N$, with $\mathbf{r}^o \geq \mathbf{r}^*$ component wise. For $\mathbf{p}$ in $\mathbb{G}_{\mathrm{p}}$, $\nexists \mathbf{p}^o = \left(p_1^o, p_2^o, ..., p_N^o\right) : u_i(\mathbf{p}^o) \geq u_i(\mathbf{p}^*), \forall i \in N$ and $u_j(\mathbf{p}^o) \geq u_j(\mathbf{p}^*)$, for some $j \in N$, with $\mathbf{p}^o \geq \mathbf{p}^*$ component wise.

As known from (18),

$$\begin{aligned} f_i(\mathbf{r}^*) &= \frac{1}{p_i}\left(\ln\left(K\gamma_i^*\right) - 1\right) \\ &= \frac{1}{p_i}\left(\ln\left(K \cdot \frac{W}{r_i^*} \cdot \frac{h_i p_i}{\sum_{j=1, j\neq i}^{N} h_j p_j + \sigma^2}\right) - 1\right) = 0. \end{aligned} \tag{2.31}$$

Without loss of generality, let $r_i^o = \rho_i r_i^*, \forall i \in N$ where $\rho_i > 1, \forall i \in N$. Then we have the following:

$$\begin{aligned} u_i^p(\mathbf{r}^o) &= r_i^o \frac{\ln\left(K\gamma_i^o\right)}{p_i} - c_i \cdot p_i \\ &= r_i^o \frac{ln\left(K \cdot \frac{W}{r_i^o} \cdot \frac{h_i p_i}{\sum_{j=1, j\neq i}^{N} h_j p_j + \sigma^2}\right)}{p_i} - c_i \cdot p_i. \end{aligned} \tag{2.32}$$

In order to find out how $u_i^p(\mathbf{r}^o)$ behaves with ρ_i, it needs to find the first order derivative of

$u_i^p(\mathbf{r}^o)$ with respect to ρ_i as follows:

$$\begin{aligned}\frac{\partial u_i^p(\mathbf{r}^o)}{\partial \rho_i} &= r_i^* \frac{1}{p_i}\left(\ln\left(K\frac{1}{\rho_i}\frac{W}{r_i^*}\cdot\frac{h_i p_i}{\sum_{j=1,j\neq i}^{N} h_j p_j+\sigma^2}\right)-1\right)\\ &= r_i^* f_i\left(\frac{1}{\rho_i}\mathbf{r}^*\right), \forall i \in N.\end{aligned} \tag{2.33}$$

It can be checked easily that $f_i\left(\frac{1}{\rho_i}\mathbf{r}^*\right) < f_i(\mathbf{r}^*) = 0, \forall i \in N$. Hence, $\frac{\partial u_i^p(\mathbf{r}^o)}{\partial \rho_i} < 0, \forall i \in N$, that is, $u_i^p(\mathbf{r}^o)$ is decreasing over $\rho_i > 1$ for all users. So we conclude that $\mathbf{r}^*$ is a Pareto-optimal Nash equilibrium point of the NPRGP game.

For transmit power part, note that

$$\begin{aligned}g_i(\mathbf{p}^*) =& r_i \cdot \frac{1-\ln\left(K\gamma_i^*\right)}{{p_i^*}^2} - c_i\\ =& r_i \cdot \frac{1-\ln\left(K\cdot\frac{W}{r_i}\cdot\frac{h_i p_i^*}{\sum_{j=1,j\neq i}^{N} h_j p_j+\sigma^2}\right)}{{p_i^*}^2} - c_i = 0.\end{aligned} \tag{2.34}$$

Without loss of generality, let $p_i^o = \delta_i p_i^*, \forall i \in N$ where $\rho_i > 1, \forall i \in N$. Then we have the following:

$$\begin{aligned}u_i^p(\mathbf{p}^o) =& r_i \frac{\ln\left(K\gamma_i^o\right)}{p_i^o} - c_i \cdot p_i^o\\ =& r_i \frac{\ln\left(K\cdot\frac{W}{r_i}\cdot\frac{h_i p_i^o}{\sum_{j=1,j\neq i}^{N} h_j p_j+\sigma^2}\right)}{p_i^o} - c_i \cdot p_i^o.\end{aligned} \tag{2.35}$$

To find the first order derivative of $u_i^p(\mathbf{p}^o)$ with respect to δ_i, we get

$$\begin{aligned}\frac{\partial u_i^p(\mathbf{p}^o)}{\partial \delta_i} =& p_i^*\left(\frac{r_i}{\left(\delta_i p_i^*\right)^2}\left(1-\ln\left(K\frac{W}{r_i}\cdot\frac{h_i\left(\delta_i p_i^*\right)}{\sum_{j=1,j\neq i}^{N} h_j p_j+\sigma^2}\right)\right)-c_i\right)\\ &+c_i \cdot p_i^*\left(\frac{1}{{\delta_i}^2}-1\right)\\ =& p_i^* g_i(\delta_i \mathbf{p}^*) + c_i \cdot p_i^*\left(\frac{1}{{\delta_i}^2}-1\right), \forall i \in N.\end{aligned} \tag{2.36}$$

Note that $g_i(\bullet)$ is a strictly decreasing function with respect to p_i. So we obtain that $g_i(\delta_i \mathbf{p}^*) < g_i(\mathbf{p}^*) = 0, \forall i \in N$. Note that $\delta_i > 1, \forall i \in N$, so $c_i \cdot p_i^*\left(\frac{1}{{\delta_i}^2}-1\right) < 0, \forall i \in N$. Hence, $\frac{\partial u_i^p(\mathbf{p}^o)}{\partial \delta_i} < 0, \forall i \in N$, that is, $u_i^p(\mathbf{p}^o)$ is decreasing over $\delta_i > 1$ for all users. Accordingly, we conclude that $\mathbf{p}^*$ it is also a Pareto-optimal Nash equilibrium point of the NPRGP game.

According to the above discussion, an asynchronous rate- and power-control algorithm is proposed which converges to a unique Nash equilibrium point $(\mathbf{r}^*, \mathbf{p}^*)$ of the game G_p.

In this algorithm, the SUs update their transmission rates and powers in an asynchronous manner that requires no strict synchronization among rate and power vectors, which is well suitable for distributed implementation. Assume that the SU i updates its rate and power at time instance in the set $T_i = \{t_{i1}, t_{i2}, ...\}$, with $t_{ik} < t_{ik+1}, t_{i0} = 0, \forall i \in N$. Let $T = \{t_1, t_2, ...\}$ where $T_i = T_1 \bigcup T_1 \bigcup ... \bigcup T_N$ with $t_k < t_{k+1}$. The NPRGP game generates a sequence of rates and powers as follows. Let ε be a small number, for example 10^{-7}. A sequence of rates and powers is obtained as follows:

Algorithm 1: (***Asynchronous Rate- and Power-Control (ARPC) Algorithm***)

Consider the non-cooperative rate- and power- control game (NPRGP) as given in (16).

1: **if** $t_{i0} = 0, \forall i \in N$ **then**

2: **set** $\mathbf{p}(0) = (p_1(0), p_2(0), ..., p_N(0))$ to any random vector $\mathbf{p}$,

set $\mathbf{r}(0) = (r_1(0), r_2(0), ..., r_N(0))$ to any random vector $\mathbf{r}$.

3: **for** all $i \in N$, such that $t_k \in T_i$, compute

$$r_i^* = \arg\max_{r_i \in R^+} u_i^p(\mathbf{p}, \mathbf{r}),$$

4: **then** set the transmission rate:

$$r_i(t_k) = \min\left(r_i^*, r^{\max}\right).$$

5: Given the prior power $p(t_{k-1})$, compute

$$p_i^* = \arg\max_{p_i \in R^+} u_i^p(\mathbf{p}, \mathbf{r}),$$

6: **then** set the transmit power:

$$p_i(t_k) = \min\left(p_i^*, p^{\max}\right).$$

7: **if** $\|p_i(t_k), p_i(t_{k-1})\| \leq \varepsilon$ and $\|r_i(t_k), r_i(t_{k-1})\| \leq \varepsilon$,

8: **stop**.

9: Declare the Nash equilibrium as $(r_i(t_k), p_i(t_k))$.

10: **else**,

k = k + 1 and go to step 2.

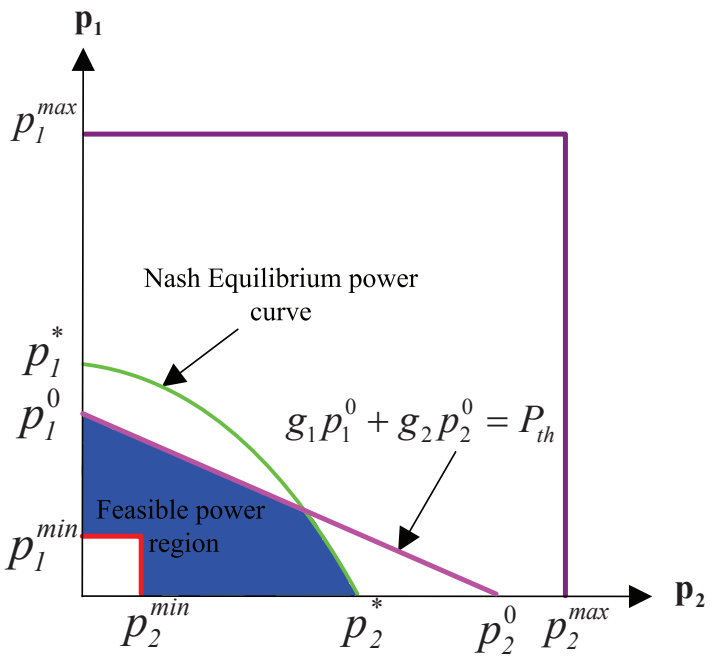

Figure 2.3. Feasible transmit power region for SUs in the CRN.

2.4 Power Pricing Factor for NPRGP

2.4.1 Feasible Power Pricing Factor

In this subsection, the process of how to find the *feasible power pricing factors* for SUs in the NPRGP that conforms the interference-power constraint (2) is discussed. To illustrate this process clearly, in this subsection, the two-SUs case is first illustrated with the referred graph, Fig.2.3 and Fig.2.4. For the multi-SUs case, it is the same idea as the two-SUs case, and the corresponding algorithm for general N-SUs is designed in subsection 2.4.1.

At first, consider an extreme case that SUs' transmit powers achieve the interference-power constraint (2). For the transmit-power vector $\{\mathbf{p}|\mathbf{p}^o = \left(p_1^o, p_2^o, ..., p_N^o\right)\}$, this is:

$$\sum_{i=1}^{N} g_i p_i^o = \mathrm{P_{th}}. \tag{2.37}$$

We call the set of power vectors that satisfy (37) as the *interference-power boundary* for SUs. Every power vector beyonds this boundary will violate (2). Geometrically, (37) represents a *super plane* in the Euclidean space $\mathbb{R}^N$ with $p_i^o, i = 1, 2, ..., N$ as variables. For the two SUs case, it is degraded to a *straight line* $g_1 p_1 + g_2 p_2 = \mathrm{P_{th}}$ in $\mathbb{R}^2$ as in Fig.2.3. The region bellow this line in the first quadrant is the feasible transmit-power region for

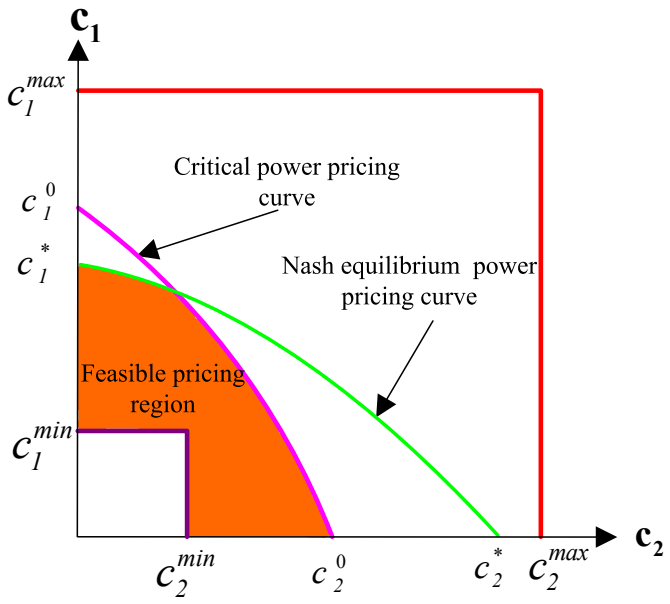

Figure 2.4. Feasible region of power pricing for SUs in the CRN.

SUs, which is the region described by inequality (2). Moreover, p_i^{max} and p_i^{min} are the *upper bound* and the *lower bound* for all SU $i = 1, 2, ..., N$ in Fig.2.3.

From theorem 3, it is known that the Nash equilibrium of transmit-power vector $\mathbf{p}^* = \left(p_1^*, p_2^*, ..., p_N^*\right)$ is related to the power-pricing factor vector $\mathbf{c}^* = \left(c_1^*, c_2^*, ..., c_N^*\right)$. Their relationship is connected by (26). By assigning different power-pricing factors, SUs will obtain different Nash equilibria of transmit power. Since increasing transmit power of one SU will also cause interference to other $N-1$ SUs and affect their utility, the Nash equilibrium point of transmit power for each SU can not exceed certain threshold, and there must be a upper boundary[1]. In Euclidean space $\mathbb{R}^2$, it is a two-dimensional curve in Fig.2.3. We call this boundary curve as *Nash equilibrium power curve*.

Until now, the *feasible power region* of SUs is obtained. The feasible power region satisfies that: each SU transmits its data on the Nash equilibrium point, while do not cause excessive interference to PUs. In geometry, this region is the intersection of Nash equilibrium, interference-power constraint, and maximum and minimum transmit-power curves in

[1]Note that the detailed algorithm that investigates this process will be available in the Numerical Results Section.

the first quadrant. The feasible power region is labeled in Fig.2.3.

Now, it is necessary to find the feasible power-pricing vectors that ensure all SUs transmit powers in the feasible power region. This process is obvious: we only need to map the power curves as in Fig.2.3 into pricing curves according to (26) and find the corresponding pricing region. For the same example in $\mathbb{R}^2$, the *feasible pricing region* is labeled in Fig.2.4. Moreover, the *critical power pricing curve* $\mathbf{c}^o = \left(c_1^o, c_2^o, ..., c_N^o\right)$ in Fig.2.4 denotes the corresponding pricing vector when transmit-power vector reaches the interference-power constraint in Fig.2.3. Similarly, we have the *Nash equilibrium pricing curve* shown in Fig.2.4. Now, we have set up the one-to-one mapping between Fig.2.3 and Fig.2.4. Mathematically, this mapping is connected by (26), where vector p_i decreases when c_i increases, and visa versa. To restrict the transmit powers of SUs that guarantee the interference-power constraint, it requires to choose high pricing factor for the SUs. Accordingly, we can seek the best interference-power pricing from the feasible pricing region (shown in Fig.2.4) that results in the optimal equilibrium rates and powers of the NPRGP. The corresponding algorithm for the general N-SUs case is designed in the following subsection.

2.4.2 Seeking for the best interference-power pricing

The equilibrium rates and powers of the NPRGP given in (16) are obtained by Algorithm 1. To obtain the best interference-power pricing for each SU that maximizes the overall networking utility, at first, we get the equilibrium rates and powers in the NPRGP with no pricing ($c_i = 0, \forall i \in N$). This is equivalent to play the NPRG given in (3). The equilibrium rates and powers in the NPRG are obtained by solving $\gamma_i = \gamma_i^{tar}$ for each user i. Once the equilibrium without pricing is obtained, the NPRG is played again after increasing the pricing factor c_i by a positive value Δc_i. Accordingly, algorithm 1 returns a set of rates and powers at the equilibrium with the incremented pricing factors. If the utilities at this new equilibrium are improved compared with previous instance and the *interference-power constraint* is conformed, the pricing factors will be increased and the procedure will be repeated again and again. Let this process continue until an increase in c_i results in

at least one SU's utility is worse than the previous equilibrium values or the *interference-power constraint* is violated. Declare the last value of c_i (*Pareto optimal* [2]) to be the best pricing factor, c_i^{best}. The process of finding c_i^{best} can be summarized as an algorithm in the following.

Algorithm 2:(***Best pricing factor for the NPRGP***)

1: **set** $c_i = 0$, announce c_i to all SUs.

2: **get** $u_i, \forall i \in N$, at equilibrium powers and rates.

3: **set** increment $c_i£ = c_i + \Delta c_i$ to satisfy *interference-power constraint*, and update c_i to SU i $\forall i \in N$.

4: **if** $u_i^{c_i} \leqslant u_i^{c_i+\Delta c_i}, \forall i \in N$.

5: **then** go to step 2.

6: **else**

7:**stop**, declare $c_i^{best} = c_i, \forall i \in N$.

The Algorithm 2 can be implemented by heuristic search, which searches one SU's pricing factor and let other ones be fixed. At the initial stage, the step size of each c_i can be set very large until step 3 and step 4 in Algorithm 2 are violated. Then, adjust each c_i with small step size to find the optimal values of the pricing factors. In the real network implementation, the information obtained in the searching processes can be accumulated to reduce the implementation complexity.

2.5 Numerical Results

2.5.1 Simulation Model

Consider a 3G cellular primary network with a radius of $3km$, where BS B_p is in the center of the circle. The secondary cell with radius $500m$ located $500m$ south (the central of the secondary cell is 1000m) to B_p as shown in Fig.2.1. There are 10 SUs in the secondary cell sharing the spread-spectrum bandwidth W, $3.84 * 10^6$ Hz (chip rate). The SUs are assumed at distances d = [50, 100, 150, 200, 250, 300, 350, 400, 450, 500] m from the

[2]Since in this case SUs' utilities can not increase cooperatively for any set of SUs.

secondary base station BS_s. The instant locations of the 10 SUs are described in Fig.2.1: SUs are distributed uniformly around BS_s, and their distances d decrease gradually in the clockwise direction. The AWGN power σ^2 at the receiver of each SUs is 10^{-15} *W/Hz*. The maximum and the minimum power-constraints for each SUs is 0.2 *Watts* and 10^{-5} *Watts*. The maximum and the minimum rate-constraints for each SU is 96 *kbps* and 0 *kbps*. There is no forward error correction, and the channel gains of secondary data communication use the simple path loss model $h_i = K/d_i^4$, where d_i is the distance between SU i and BS_s. For the interference path-loss model g_i from SU i to BS_p, we use the same model as h_i. PUs are randomly located around the B_p, and their tolerable interference power at receiver side (B_p in Fig.2.1) is assumed to be $-118dBm$. The noise power spectrum density at the receiver side of BS_p, same as SU, is $-120dBm$. We consider the required target SINR γ_i^{tar} for each SU to be 12.42 (equal to the equilibrium SINR obtained by the algorithm in [42]). Then we calculate the value of parameter 'K' using (9), where $K = 0.21886$.

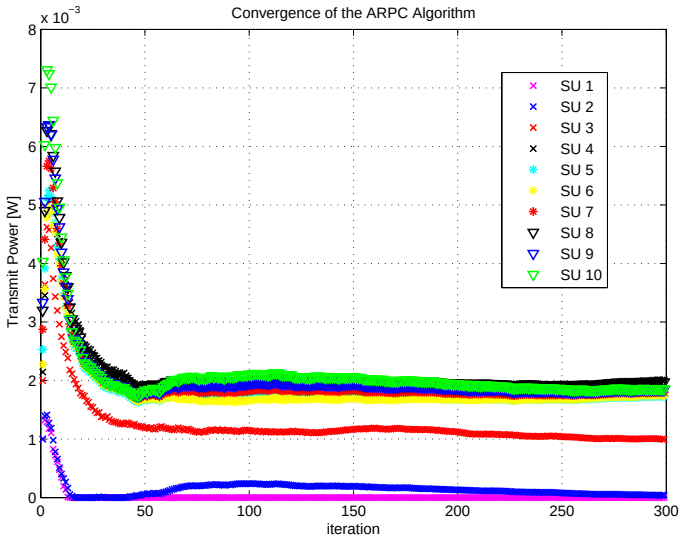

Figure 2.5. Convergence of transmit powers of all SUs.

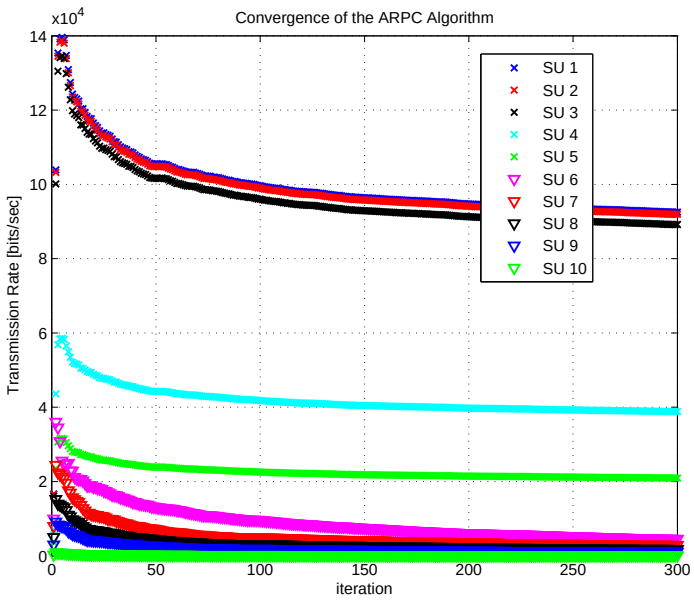

Figure 2.6. Convergence of transmission rates of all SUs.

2.5.2 Performance Analysis

The algorithms 1 and 2 are implemented to find the Nash Equilibrium of transmit power, transmission rate and best pricing factor for each SU, the corresponding simulations are shown in Fig.2.5-2.10. The algorithms are implemented for the following games: non-cooperative joint-power and rate-control game with pricing (NPRGP), non-cooperative joint-power and rate-control game (NPRG), non-cooperative power-control game with pricing (NPGP). The NPRGP is the central problem of this paper, which is formulated in (16) and fully implemented by algorithm 1 and 2. The NPRG is the problem formulated in (3), where pricing for interference power is omitted. For the algorithmic implementation of the NPRG, we only need to set all the $c_i, \forall i \in N$, equal to zero. The NPGP is the conventional power control problem in CRNs under interference-power constraint, like [10][11][13-18], where only power control is considered. In this case, the transmission rate for each SU is fixed, and it is the average value of transmission rates obtained in the NPRGP (It is shown in Fig. 2.9). For the algorithmic implementation of the NPGP, we only need to omit the transmission-rate control in algorithm 1, and set them fixed values; the algorithm 2 for the

NPGP is the same.

In Fig.2.5, it is found that the transmit powers of SUs take around 300 iterations to converge when we set $\varepsilon = 10^{-7}$. In our simulation, a typical iteration of algorithm 1 need time $\Theta(1ns)$[3]. The total convergence time for algorithm 1 is $\Theta(1\mu s)$. For the message exchange and utility comparison between SUs, in our experiments, the time consumed based on the heuristic-search approach, on the average, is $\Theta(400ms)$. We believe that the implementation delay of these algorithms is tolerable for a typical cellular communication cycle [69]. In Fig.2.5, it is found that the transmission rates converge more "smooth" and much "faster" than transmit powers in Fig.2.5. This is because there are no extra constraints that affect the convergence of transmission rates.

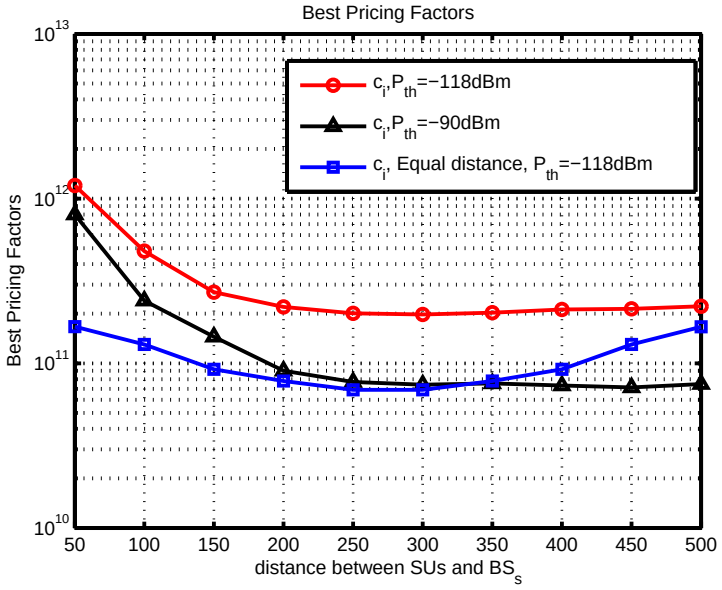

Figure 2.7. Best pricing factors of SUs by the algorithm 2.

In Fig.2.7, the best pricing factors obtained by algorithm 2 under interference-power constraints $P_{th} = -118$dBm and $P_{th} = -90$dBm are shown. In the light interference-power constraint condition $P_{th} = -118$dBm, it is observed that the nearer the SU to the base station BS_s, the larger the pricing factor. At this time, the impact of inter-SUs interference on SUs'

[3]Here we use the Knuth's notation "the big-theta" [68] in algorithm analysis. $f(x) = \Theta(g(x))$ means you can say that $f(x)$ is the same order of $g(x)$.

utilities is the leading part, compared with the impact of interference-power constraint from PUs. Moreover, when a SU is near to BS_s, more interference will be caused to other SUs, so the penalty of the interference power on this SU will be larger. This means the price for the SU's data transmission will be larger. In the strong interference-power constraint condition $P_{th} = -90\text{dBm}$, it is observed that the pricing factors are larger, when SUs are more near to PUs. At this time, SUs are impacted a lot by PUs' interference-power constraint besides the inter-SUs interference. Moreover, the pricing factors are investigated in the special network scenario where the SUs are of equal distance $50m$ to the secondary base station BS_s. To ease the comparison, the results are shown in Fig.2.7, where now the distance between SUs and BS_s are $d = [50, 50, 50, 50, 50, 50, 50, 50, 50, 50]m$. By simply checking these pricing factors, it is found that their difference approximately obey the rule in (38). At this time, the utility difference among SUs are only affected by the difference of their channel gains to PUs.

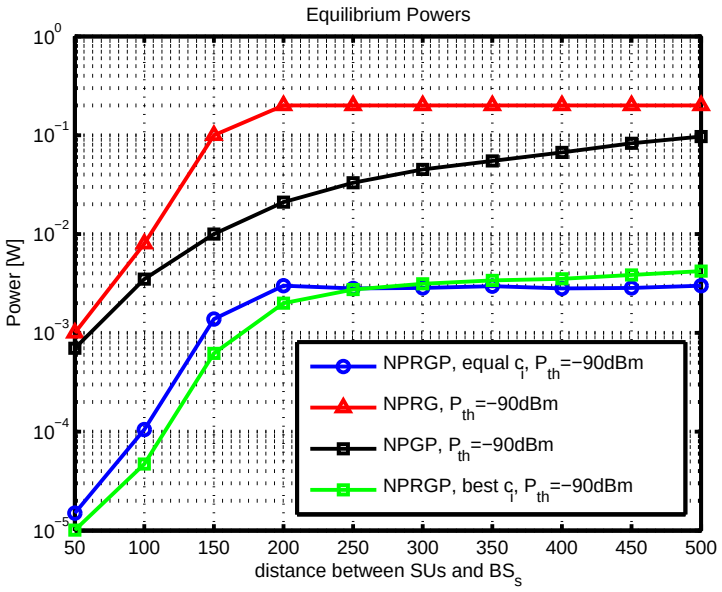

Figure 2.8. Comparison of Nash equilibrium of transmit powers.

In Fig.2.8, the Nash equilibrium powers of SUs obtained by algorithm 1 and 2 are compared. At First, with nearly the same transmission rate for each SU shown in Fig.2.9, the transmit power for the NPRGP is greatly reduced when compared with the NPRG in

(3) without price. This conforms our statement that price is an effective tool to reduce interference and lead the system to work in an efficient operating point. Second, when considering the conventional power-control problem for SUs with interference-power pricing, the NPGP, it is observed that the transmit powers for SUs systems are larger than the transmit powers in the NPRGP. This means that under the same transmission rates for all SUs, the NPRGP use less transmit powers than the NPGP, which conforms the central point in this research: joint-power and rate-control can achieve better radio-resource utilization than the only power control in CRNs.

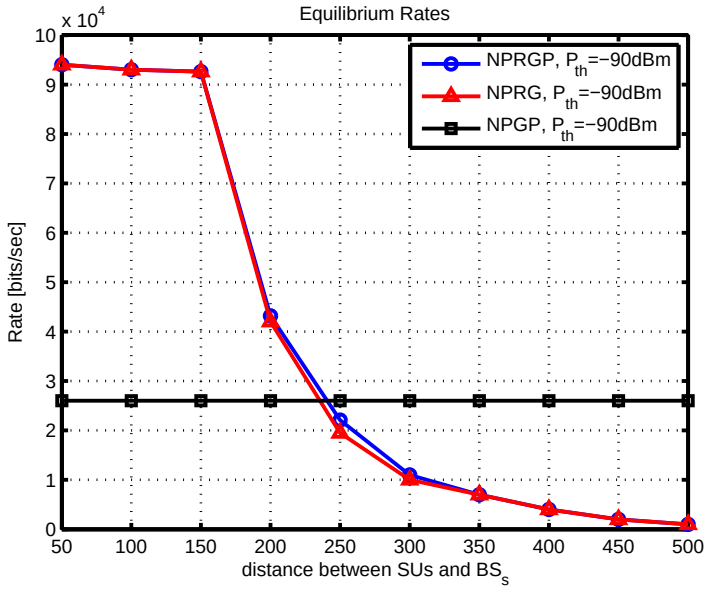

Figure 2.9. Comparison of Nash equilibrium of transmission rates.

In Fig.2.9, the Nash equilibrium rates of SUs are compared. The obtained transmission rate of each SU for the NPRGP and the NPRG is nearly the same. In these games, each SU is allocated optimal transmission rate according to the distance between the SU and BS_p. For the NPGP, every SU is allocated the same transmission rate to guarantee its target SINR. The overall performance for SUs in the NPGP is inferior to the NPRGP and the NPRG. Notice that the transmission rates of the NPGP are greater than that of the NPRGP and the NPRG, when the distance between SUs and BS_p is greater than $250m$. However, the corresponding transmit power for each SU, in Fig.2.8, is largely increased.

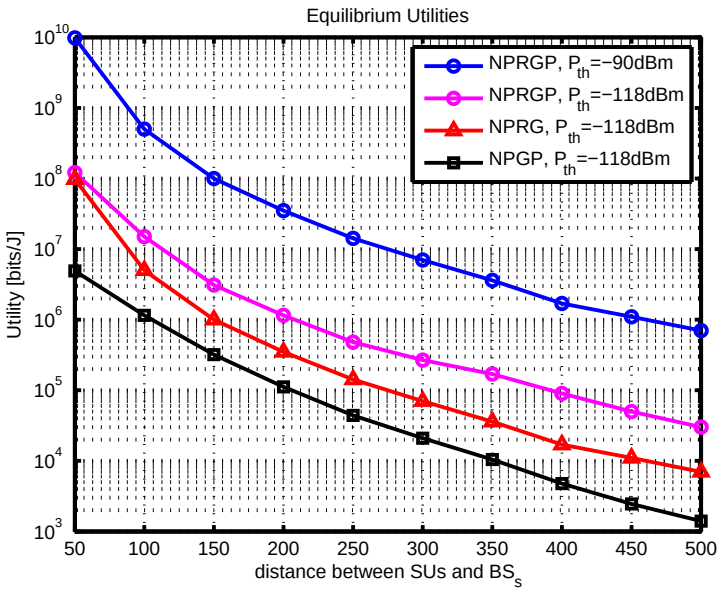

Figure 2.10. Comparison of SUs' utilities.

In Fig.2.10, the utilities or satisfactions experienced by SUs of the proposed games, i.e., NPRGP, NPRG and NPGP, are compared. As indicated in Fig.2.8, by introducing price into the NPRGP, the transmit powers can be largely reduced when achieving nearly the same transmission rates as in the NPRG. in Fig.2.10, with the best pricing factor for each SU, the utility of each SU can be greatly improved. Moreover, the utilities of the NPRGP are computed under the strong interference-power constraint condition $P_{th} = -118$dBm. It is observed that the utilities are less than the utilities in the light interference-power constraint condition $P_{th} = -90$dBm, and the SUs that are near to PUs suffer more utilities decrease than the SUs that are far from PUs.

CHAPTER 3

ASYNCHRONOUS POWER-CONTROL GAME WITH CHANNEL OUTAGE CONSTRAINTS IN COGNITIVE RADIO NETWORKS

3.1 System Model

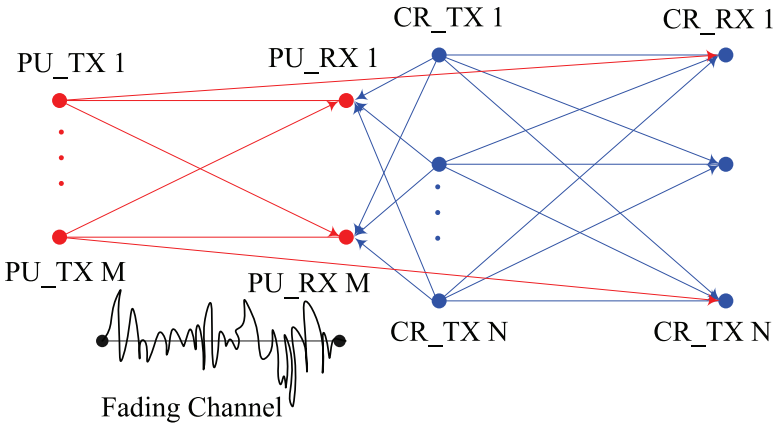

Figure 3.1. *N* **CR users and** *M* **PUs sharing a single-channel case.**

Consider the scenario shown in Fig.3.1, in which there are a set of $\mathcal{N} := \{1, ..., N\}$ CR transceiver pairs allowed to share a set of $\mathcal{K} := \{1, ..., K\}$ channels used for a primary network. There are $\mathcal{M} := \{1, ..., M\}$ primary transceiver pairs co-located with the CR users having priority in using these K channels. The behavior of the primary and CRNs are as follows: CR users are playing the power-control game under the influence from primary system; PU pairs transmit with fixed power for the duration of this game, where their transmission are suffered by CR users' transmit power and dynamic fading environments. Thus, we can denote the transmit power of CR user n and PU m on channel k as p_n^k and $\bar{p}_m^k$, respectively. We assume the Rayleigh fading in the model. Let g_{ij}^k denote the slow-varying channel gain between i's CR or PU transmitter and j's CR or PU receiver, which is assumed to be stationary and do not change significantly over the time scale of this analysis. Let f_{ij}^k denote the fast time-scale Rayleigh-fading-channel gain between i's CR or PU transmitter and j's CR or PU receiver, which is assumed to be a "block" fading model [69] (static over individual data frames but varying from one frame to another) with unit mean independent

exponentially distributed random variable. The received SINR of any CR user or PU l on channel k is given by

$$\gamma_l^k(\mathbf{p}^k) = \frac{g_{ll}^k f_{ll}^k p_l^k}{\sum_{i\neq l}^{N} g_{il}^k f_{il}^k p_i^k + \sum_{j\neq l}^{M} g_{jl}^k f_{jl}^k \bar{p}_j^k + \sigma_k^2}, \tag{3.1}$$

where σ_k^2 is the noise power, and $\mathbf{p}^k = \{p_n^k; \forall n \in \mathcal{N}\}$ is the power vector of all CR users' transmit powers on channel k. Note that $\{\bar{p}_m^k, \forall m \in \mathcal{M}\}$ is a constant value. In a Rayleigh-fading environment, $\gamma_l^k(\mathbf{p}^k)$ is a random variable with a very complex distribution [70], since it is the ratio of an exponential random variable to a sum of exponential random variables. The careful choice of $\bar{\gamma}_l^k$ establishes the intended QoS request of each user on a specific channel. The QoS requested is provided when the SINR exceeds a given threshold $\bar{\gamma}_l^k$. For the flat block-fading channels, the proportion of time that certain target SINR threshold $\bar{\gamma}_l^k$ is not satisfied is defined as the outage probability of user l on channel k [69]. The outage probability $Q_l^k = \Pr\left(\gamma_l^k \leq \bar{\gamma}_l^k\right)$ of l^{th} user (CR users or PUs) on channel k can be found in [70] of a closed form expression:

$$Q_l^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) = 1 - \exp(-\tfrac{\sigma_k^2 \bar{\gamma}_l^k}{g_{ll}^k p_l^k}) \prod_{i\neq l}^{N} (1 + \tfrac{\bar{\gamma}_l^k g_{il}^k p_i^k}{g_{nn}^k p_n^k})^{-1} \bullet \prod_{j\neq l}^{M} (1 + \tfrac{\bar{\gamma}_l^k g_{jl}^k \bar{p}_j^k}{g_{nn}^k p_n^k})^{-1}, \tag{3.2}$$

where $\mathbf{p}_{-n}^k = (\mathrm{p}_1^k, \mathrm{p}_2^k, ..., \mathrm{p}_{n-1}^k, \mathrm{p}_{n+1}^k, ..., \mathrm{p}_N^k)$. Denote $\zeta_l^k = 1 - \exp(-\frac{\sigma_k^2 \bar{\gamma}_1^k}{G_{11}^k p_1^k})$. It is the outage probability of primary receiver in Rayleigh-fading channel with exponential distribution f_l^k and unit mean in the absence of other PUs or CR users. Therefore, the above equation indicates the outage probability of a single user in the present of multiple PUs and/or CR users. Accordingly, the aim of each CR user n is to transmit with the target data rate $\bar{R}_n^k$ (it is one-to-one correspondent to $\bar{\gamma}_n^k$ under any fixed modulated schemes or under the information-theoretical capacity scheme) on channel k with the specified channel outage ξ_n^k constraint by using as little power consumption as possible,

$$Q_n^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_n^k, \forall n \in \mathcal{N}. \tag{3.3}$$

Inequality (3.3) denotes the outage constraint of the CR user n on the channel k that guarantees its requested QoS. Similarly, we have the definition for each PU's target data rate $\bar{R}_m^k$,

and the outage constraints of each PU m on each channel k, which requests transmit-power restriction on each CR user.

$$Q_m^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_m^k, \forall m \in \mathcal{M}. \tag{3.4}$$

Notice that there exists the co-channel interference among CR users and PUs. The "coupled" effects among them impact the QoS performance of all PUs and CR users. Therefore, this coupled power-control problem can be modulated as a non-cooperative game as follows:

$$\mathrm{G}_{MPCF} = \{\mathcal{N}, \{\Psi_n\}_{n\in\mathcal{N}}, \{\mathcal{X}_n\}_{n\in\mathcal{N}}\}. \tag{3.5}$$

The cost function of CR user n's is the total power allocated over all channels:

$$\Psi_n(\mathbf{p}_n) = \alpha_n \sum\nolimits_{k\in K} p_n^k, \tag{3.6}$$

where α_n is the weighting factor on power-allocation vector $\mathbf{p}_n = (\mathrm{p}_n^1, \mathrm{p}_n^2, ..., \mathrm{p}_n^K)$ over all the K channels. $\mathcal{X}_n$ is each CR user's strategy that are coupled with other users' strategies and can be expressed as:

$$\mathcal{X}_n(\mathbf{p}_{-n}) = \{\mathbf{p}_n \in \Omega_n^P | Q_n^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_n^k;\ Q_m^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_m^k, \forall n \in \mathcal{N}, \forall m \in \mathcal{M}\} \tag{3.7}$$

where $\Omega_n^P = \{\mathrm{p}_i \in R^K | \sum_{k\in K} \mathrm{p}_n^k \leq P_n^{\max}; pl_n^k \leq \mathrm{p}_n^k \leq pu_n^k, \forall k \in K, \forall n \in N\}$ is CR user n's transmit-power range, and $P_n^{\max}$ is the total power budget of CR user n. pu_n^k and pl_n^k denote the upper and lower bound of CR user n's transmit power on channel k. The power vector $\mathbf{p}_{-\mathrm{n}} = (\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_{n-1}, \mathbf{p}_{n+1}, ..., , \mathbf{p}_N)$ denotes the transmit power of all the other CR users. From (3.7), the whole strategy space of the G_{MPCF} can be expressed as:

$$\begin{aligned} X = \{&(\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N) \in \Omega_1^P \times, ..., \times \Omega_{\mathrm{N}}^P | Q_n^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_n^k; \\ &Q_m^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_m^k, \forall m \in \mathcal{M}, \forall k \in \mathcal{K}.\} \end{aligned} \tag{3.8}$$

Given $\mathbf{p}_{-n}$, each CR user $n \in \mathcal{N}$ are playing the multi-channel power-control game with the following objectives:

$$\min_{\mathbf{p}_n \in X_n(\mathbf{p}_{-n})} \Psi_n(\mathbf{p}_n) \tag{3.9}$$

Based on the definition of Nash equilibrium [71], the equilibrium power vector $\mathbf{p}_{\mathrm{n}} = (\mathbf{p}_1^*, \mathbf{p}_2^*, ..., \mathbf{p}_N^*)$ is obtained by the following equation:

$$\mathbf{p}_n^* = \arg \min_{\mathbf{p}_n \in \mathcal{X}_n(\mathbf{p}_{-n})} \Psi_n(\mathbf{p}_n), \forall n \in \mathcal{N} \tag{3.10}$$

Definition 1 A target QoS profile for both CR users and PUs $(\{\bar{\gamma}_l^k(\mathbf{p}^k\}, \xi_n^k, \xi_m^k), \forall n \in \mathcal{N}, \forall m \in \mathcal{M}$ is *feasible*, if there exist a transmit-power vector $(\mathbf{p}_1, ..., \mathbf{p}_N)$ such that $X = \{(\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_N) \in \Omega_1^P \times, ..., \times \Omega_{\mathrm{N}}^P | Q_n^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_n^k; Q_m^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_m^k, \forall m \in \mathcal{M}, \forall k \in \mathcal{K}\}$.

3.2 Properties of the Power-Control Game

In this section, the properties of the proposed power-control game are investigated. To view this problem clearly, we may rewrite the game-theory problem to an optimization format as follows.

$$\begin{aligned} \min_{\mathbf{p}_n \in X_n(\mathbf{p}_{-n})} \ & \Psi_n(\mathbf{p}_n), \forall n \in \mathcal{N} \\ s.t. \quad & Q_n^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_n^k; \\ & Q_m^k(\mathrm{p}_n^k, \mathbf{p}_{-n}^k) \leq \xi_m^k; \\ & \mathbf{p}_n \in \Omega^P; \\ & \forall m \in \mathcal{M}, \forall k \in \mathcal{K}. \end{aligned} \tag{3.11}$$

After expanding the first constraint of problem (3.11), we obtain

$$1 - \xi_n^k \leq \exp\left(\frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k p_n^k}\right) \prod_{i \neq n}^{N} \left(1 + \frac{\bar{\gamma}_n^k g_{in}^k p_i^k}{g_{nn}^k p_n^k}\right)^{-1} \prod_{j=1}^{M} \left(1 + \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k}{g_{nn}^k p_n^k}\right)^{-1}. \tag{3.12}$$

Similarly, the second constraint of the problem (3.11) is expanded as

$$\prod_{n=1}^{N} (1 + b_{nm}^k p_n^k) \leq \frac{a_m^k}{1 - \xi_m^k}, \tag{3.13}$$

where $a_m^k = \exp\left(\frac{\sigma_k^2 \bar{\gamma}_m^k}{g_{mm}^k \bar{p}_m^k}\right) \prod_{j \neq m}^{M} \left(1 + \frac{\bar{\gamma}_m^k g_{jm}^k \bar{p}_j^k}{g_{mm}^k \bar{p}_m^k}\right)^{-1}$ and $b_{nm}^k = \frac{\bar{\gamma}_m^k g_{nm}^k}{g_{mm}^k \bar{p}_m^k}$ are constant values.

The feasible power vectors in the constraints (3.12) and (3.13) are non-convex. Thus, the constrained game (3.11) is nonlinear and non-convex. The geometric programming [60] can be applied to solve this problem (3.11). However, generally geometric programs are not convex optimization problems, and they can be transformed into convex problems by

adopting logarithmic change of variables: let $\tilde{p}_n^k = \log(p_n^k)$. Therefore, the transformed problem of (3.11) is given by

$$\begin{aligned}
&\min_{\tilde{\mathbf{p}}_n \in X_n(\tilde{\mathbf{p}}_{-n})} \ \Psi_n(\mathrm{e}^{\tilde{\mathbf{p}}_n}), \forall n \in \mathcal{N} \\
&\quad s.t. \ \ \sum_{i\neq n}^{N} \log(1+\frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) + \sum_{j=1}^{M} \log(1+\frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) + \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}} + \log(1-\xi_n^k) \leq 0; \\
&\qquad \sum_{n=1}^{N} \log(1+b_{nm}^k \mathrm{e}^{\tilde{p}_n^k}) - \log\frac{a_m^k}{1-\xi_m^k} \leq 0; \\
&\qquad \tilde{\mathbf{p}}_n \in \Omega^{\tilde{P}}; \forall m \in \mathcal{M}, \forall k \in \mathcal{K}.
\end{aligned} \tag{3.14}$$

Now, we investigate the properties of problem (3.14). We denote $\tilde{Q}_n^k(\tilde{\mathrm{p}}_n^k, \tilde{\mathbf{p}}_{-n}^k)$ and $\tilde{Q}_n^k(\tilde{\mathrm{p}}_n^k, \tilde{\mathbf{p}}_{-n}^k)$ as the function notations of the first and second constraints on CR users and PUs in problem (3.14). The following two lemmas indicate the properties of the constraints of CR users and PUs in (3.14), respectively.

Lemma 1 The CR user outage-constraint function $\tilde{Q}_n^k(\tilde{\mathrm{p}}_n^k, \tilde{\mathbf{p}}_{-n}^k)$ in (3.14) have the following properties.

- **P.1** $\tilde{Q}_n^k$ is *strictly* decreasing in $\tilde{p}_n^k$
- **P.2** $\tilde{Q}_n^k$ is *convex* in $\tilde{p}_n^k$
- **P.3** $\tilde{Q}_n^k$ is strictly *increasing* and *convex* in $\tilde{p}_i^k, \forall i \neq n, i \in \mathcal{N}$
- **P.4** The Hessian matrix of $\tilde{Q}_n^k$ with respect to power vector $\mathbf{p}^k$ is *positively* defined.

Proof To prove the properties of **P.1** $\sim$ **P.4**, it only needs to check the first and second derivatives of $\tilde{Q}_n^k$ with respect to $\tilde{p}_n^k$ and/or $\tilde{p}_i^k$. The **P.1**, **P.2**, and **P.3** is indicated by the following 4 inequations.

$$\frac{\partial \tilde{Q}_n^k}{\partial \tilde{p}_n^k} = \sum_{i\neq n}^{N} \frac{-\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}} + \sum_{j=1}^{M} \frac{-\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k \bar{p}_j^k} - \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}} < 0. \tag{3.15}$$

$$\frac{\partial^2 \tilde{Q}_n^k}{\partial \tilde{p}_n^{k^2}} = \sum_{i\neq n}^{N} \frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k} * g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}{\left(g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}\right)^2} + \sum_{j=1}^{M} \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k * g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}{\left(g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k\right)^2} + \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}} > 0. \tag{3.16}$$

$$\frac{\partial \tilde{Q}_n^k}{\partial \tilde{p}_i^k} = \frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}} > 0. \tag{3.17}$$

$$\frac{\partial^2 \tilde{Q}_n^k}{\partial \tilde{p}_i^{k^2}} = \frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k} * g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}{\left(g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}\right)^2} > 0. \tag{3.18}$$

To prove the **P.4** for the function $\tilde{Q}_n\left(\tilde{\mathbf{p}}^k\right)$, at first, we know

$$\frac{\partial^2 \tilde{Q}_n^k}{\partial \tilde{p}_n^k \partial \tilde{p}_i^k} = \frac{-\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k} * g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}{\left(g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}\right)^2} < 0.$$

To show the Hessian Matrix $\mathbf{H}_{\tilde{Q}_n}$ is positive definite: for all vectors v,

$$\begin{aligned}
&\mathrm{v}^T \mathbf{H}_{\tilde{Q}_n} \mathrm{v} = \sum_{n=1}^{N} \mathrm{v}_n^2 \frac{\partial \tilde{Q}_n^k}{\partial \tilde{p}_n^{k^2}} + \sum_{n=1}^{N} \sum_{i \neq n}^{N} \mathrm{v}_n \mathrm{v}_i \frac{\partial^2 \tilde{Q}_n^k}{\partial \tilde{p}_n^k \partial \tilde{p}_i^k} \\
&= \sum_{n=1}^{N} \mathrm{v}_n^2 \frac{\partial \tilde{Q}_n^k}{\partial \tilde{p}_n^{k^2}} - \sum_{n=1}^{N} \sum_{i \neq n}^{N} \mathrm{v}_n \mathrm{v}_i \left| \frac{\partial^2 \tilde{Q}_n^k}{\partial \tilde{p}_n^k \partial \tilde{p}_i^k} \right| \\
&\geq \sum_{n=1}^{N} \mathrm{v}_n^2 \frac{\partial \tilde{Q}_n^k}{\partial \tilde{p}_n^{k^2}} - \sum_{n=1}^{N} \sum_{\substack{i=1 \\ i \neq n}}^{N} \mathrm{v}_i^2 \left| \frac{\partial^2 \tilde{Q}_n^k}{\partial \tilde{p}_n^k \partial \tilde{p}_i^k} \right| \\
&= \sum_{n=1}^{N} \mathrm{v}_n^2 \left(\begin{array}{l} \sum_{i \neq n}^{N} \frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k} * g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}{\left(g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}\right)^2} + \sum_{j=1}^{M} \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k * g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}{\left(g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k\right)^2} \\ + \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}} - \sum_{i \neq n}^{N} \frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k} * g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}{\left(g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}\right)^2} \end{array} \right) \\
&= \sum_{n=1}^{N} \mathrm{v}_n^2 \left(\sum_{j=1}^{M} \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k * g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}{\left(g_{nn}^k \mathrm{e}^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k\right)^2} + \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}} \right) > 0.
\end{aligned} \tag{3.19}$$

For the first inequality in the above equation, it is because of the Cauchy-Schwarz inequality: $(a^T a)(b^T b) \geq (a^T b)^2$ where $a_n = \sum_{n=1}^{N} \sum_{i \neq n}^{N} \mathrm{v}_i \sqrt{\frac{\partial^2 \tilde{Q}_n}{\partial \tilde{p}_i^k \partial \tilde{p}_n^k}}$ and $b_{nm} = \sum_{n=1}^{N} \sum_{j \neq n}^{N} \mathrm{v}_j \sqrt{\frac{\partial^2 \tilde{Q}_n}{\partial \tilde{p}_n^k \partial \tilde{p}_j^k}}$. Therefore, $\tilde{Q}_n\left(\tilde{\mathbf{p}}^k\right)$ is a strictly convex function of vector $\tilde{\mathbf{p}}^k$, and its Hessian is positive definite.

Lemma 2 The PU outage-constraint function $\tilde{Q}_m^k(\tilde{\mathrm{p}}_m^k, \tilde{\mathbf{p}}_{-n}^k)$ in (3.14) have the following properties.

- **P.1** $\tilde{Q}_m^k$ is strictly *increasing* in $\tilde{p}_n^k$
- **P.2** $\tilde{Q}_m^k$ is *convex* in $\tilde{p}_n^k$

- **P.3** The Hessian matrix of $\tilde{Q}_m^k$ with respect to power vector $\mathbf{p}^k$ is a *diagonal* matrix and is *positively* defined.

Proof Similarly, to prove the properties of **P.1** $\sim$ **P.3**, we only need to check the first and second derivatives of $\tilde{Q}_m^k$ with respect to $\tilde{p}_n^k$ and/or $\tilde{p}_i^k$. The **P.1** and **P.2** are indicated by the following 2 inequations.

$\frac{\partial \tilde{Q}_m^k}{\partial \tilde{p}_n^k} = \frac{b_{nm}^k e^{\tilde{p}_n^k}}{1+b_{nm}^k e^{\tilde{p}_n^k}} > 0;$ $\frac{\partial^2 \tilde{Q}_m^k}{\partial \tilde{p}_n^{k^2}} = \frac{b_{nm}^k e^{\tilde{p}_n^k}}{(1+b_{nm}^k e^{\tilde{p}_n^k})^2} > 0.$

To prove the **P.3**, it is easy to know the Hessian matrix $\mathbf{H}_{\tilde{Q}_m}$ of the function $\tilde{Q}_m\left(\tilde{\mathbf{p}}^k\right)$ is a diagonal matrix of the following form:

$$\mathbf{H}_{\tilde{Q}_m} = diag(c_1, ..., c_n, ..., c_N),$$

where $c_n = \frac{b_{nm}^k e^{\tilde{p}_n^k}}{(1+b_{nm}^k e^{\tilde{p}_n^k})^2}$.

For vector v, $\mathrm{v}^T H_{\tilde{Q}_m} \mathrm{v} = \mathrm{diag}(c_1 v_1^2, ..., c_N v_N^2)$. Obviously, matrix $H_{\tilde{Q}_m}$ is positive define and $\tilde{Q}_m\left(\tilde{\mathbf{p}}^k\right)$ is a strictly convex function of vector $\tilde{\mathbf{p}}^k$.

Theorem 1 *[Uniqueness]* If the target QoS profile is feasible, the transformed power-control game in (3.14) admits a unique inner Nash equilibrium solution.

Proof The proof of this theorem is similar to the one of Theorem 2 in [26]. It is briefly outlined here for completeness. Let the $\mathbf{p}_k = (\mathbf{p}_1, ..., \mathbf{p}_N)$ stratifies the feasibility of target QoS. Clearly, $\mathbf{p}_k$ is closed and bounded on Ω^P, hence it is compact. Furthermore, according to lemma 1 and 2, the Hessian matrix of constraints $\tilde{Q}_n^k$ and $\tilde{Q}_m^k$ are all positive defined with respect to power vector $\mathbf{p}_k$, which indicate strict convexity of the constrained game (3.14). By a standard theorem of game theory (Theorem 4.4 p.176 in [71]), the power-control game admits a unique Nash equilibrium.

3.3 Dual Decomposition with Layered Structures

The optimization problem in (3.14) is a game-theoretical problem with coupled constraints. Thus, distributive algorithms that converge to the N.E for the G_{MPCF} is difficult, since the

satisfaction of both QoS of CR users and PUs constraints requires the coordination among CR users. Therefore, a partial dual decomposition method for constrained game is used to relax the constraints [67]. The corresponding Lagrangian function can be given as follows:

$$\begin{aligned} \mathrm{L}_n(\tilde{\mathbf{p}}_n,\tilde{\mathbf{p}}_{-\mathrm{n}},\lambda_n,\mu) = \alpha_n \sum_{k=1}^{K} \mathrm{e}^{\tilde{p}_n^k} + \sum_{k=1}^{\mathrm{K}} \lambda_n \sum_{i\neq n}^{N} \log(1+\frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) \\ + \sum_{j=1}^{M} \log(1+\frac{\bar{\gamma}_n^k g_{jn}^k \tilde{p}_j^k}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) + \log(1-\xi_n^k) + \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}\} \\ + \sum_{m=1}^{M}\sum_{k=1}^{K} \mu_m^k \{\sum_{n=1}^{N} (\log(1+b_{nm}^k \mathrm{e}^{\tilde{p}_n^k}) - \log\frac{a_m^k}{1-\xi_m^k}\}, \end{aligned}$$

where λ_n denotes the dual price for the requested QoS of CR user n. Define the $1 \times N$ dual price vector $\boldsymbol{\lambda}_N := \{\lambda_1, ..., \lambda_N\}$. For the PUs system, define the $1 \times MK$ dual vector $\boldsymbol{\mu}_M := \{\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_M\}$, where each $1 \times K$ vector $\boldsymbol{\mu}_m := \{\mu_m^1, ..., \mu_m^k\}, \forall m \in \mathcal{M}$. It represents the set of dual prices for the PU m suffering from the interference power of the CR users over the K channels, which are known to all the CR users at the same time. The following function defines the parameterized power control objective function for $\forall n \in \mathcal{N}$

$$\begin{aligned} \Pi_n(\tilde{\mathbf{p}}_n, \tilde{\mathbf{p}}_{-\mathrm{n}}, \lambda_n, \mu) = \sum_{k=1}^{K} \{\alpha_n \mathrm{e}^{\tilde{p}_n^k} + \sum_{m=1}^{M} \mu_m^k (\log(1+ b_{nm}^k \mathrm{e}^{\tilde{p}_n^k})\} \\ + \lambda_n \sum_{k=1}^{\mathrm{K}} \{\sum_{i\neq n}^{N} \log(1+\frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) + \sum_{j=1}^{M} \log(1+\frac{\bar{\gamma}_n^k g_{jn}^k \mathrm{e}^{\tilde{p}_j^k}}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) + \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}\}. \end{aligned} \tag{3.20}$$

In addition, the coupled strategy game G_{MPCF} can be vertically decomposed into three subproblems as follows:

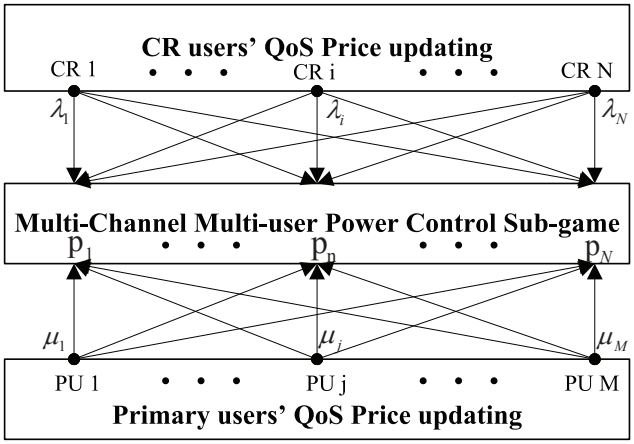

Figure 3.2. Dual decomposition with layered structures.

1) The multi-channel power control sub-game among all the CR users under the QoS

requirement of both CR users and PUs, given that the dual price vector λ and μ:

$$\mathrm{G}_{MPCF\text{-}S} = \{\mathcal{N}, \{s_n\}_{n\in\mathcal{N}}, \{\Pi_n(\tilde{\mathbf{p}}_n, \tilde{\mathbf{p}}_{-\mathrm{n}}, \lambda_n, \mu)\}_{n\in\mathcal{N}}\} \tag{3.21}$$

2) The sub-gradient updating of each CR user's dual price for the requested QoS given the equilibrium transmit-power vectors $(\mathbf{p}_1^*(\lambda,\mu), ..., \mathbf{p}_N^*(\lambda,\mu))$ for the sub-game $\mathrm{G}_{MPCF\text{-}S}(\lambda,\mu)$:

$$\lambda_n = \lambda_n + \varsigma_1 \sum_{k=1}^{\mathrm{K}} \{\sum_{i\neq n}^{N} \log(1 + \frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) + \sum_{j=1}^{M} \log(1 + \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) + \log(1-\xi_n^k) + \frac{\gamma_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}\} \tag{3.22}$$

3) The sub-gradient updating of each PU's dual price for the requested QoS given the interference equilibrium transmit-power vectors $(\mathbf{p}_1^*(\lambda,\mu), ..., \mathbf{p}_N^*(\lambda,\mu))$ for the sub-game $\mathrm{G}_{MPCF\text{-}S}(\lambda,\mu)$:

$$\mu_m^k = \mu_m^k + \varsigma_2 \{\sum_{n=1}^{N} (\log(1 + b_{nm}^k \mathrm{e}^{\tilde{p}_n^k}) - \log\frac{a_m^k}{1-\xi_m^k}\}, \forall m, k, \tag{3.23}$$

where ς_1 and ς_2 denote the updating step sizes.

Note that the sub-problems 2) 3) are maximizing the dual function

$$\mathrm{D}(\lambda,\mu) = \sum\nolimits_{n\in N} \mathrm{L}_n(\mathbf{p}_n^*(\lambda,\mu), \mathbf{p}_{-n}^*(\lambda,\mu), \lambda_n, \mu) \tag{3.24}$$

by using the sub-gradient method. Since the sub-gradient updating of λ_n and μ_m^k are adjusted from zero, so the sub-gradient of $\mathrm{D}(\lambda,\mu)$ at $\lambda_n, \forall n \in \mathcal{N}$, and $\mu_m^k, \forall m \in \mathcal{M}, k \in \mathcal{K}$ are given by;

$$\varpi_{\mathrm{k}} = \sum_{i\neq n}^{N} \log(1 + \frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\tilde{p}_i^k}}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) + \sum_{j=1}^{M} \log(1 + \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}) + \log(1-\xi_n^k) + \frac{\gamma_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\tilde{p}_n^k}}, \forall k \in \mathcal{K}.$$

$$\varpi_m^k = \sum_{n=1}^{N} (\log(1 + b_{nm}^k \mathrm{e}^{\tilde{p}_n^k}) - \log\frac{a_m^k}{1-\xi_m^k}, \forall m \in \mathcal{M}, k \in \mathcal{K}.$$

The layered structure is of a Stackelberg model as shown in Fig.3.2: each CR user n announces its QoS outage constraint dual price λ_n and each PU m announces its QoS outage constraint dual price μ_m^k to the multichannel power-control subproblem $\mathrm{G}_{MPCF\text{-}S}$ among all CR users.

From theorem 1, the primary problem in (3.14) is convex, and the sub-gradient approaches can be always used without having performance loss. Different from G_{MPCF}, the

sub-game $G_{MPCF\text{-}S}$ is only played within each CR user's power budget range $\tilde{\mathbf{p}}_n \in \Omega^{\tilde{P}}$, and the power-control problem can be represented as follows:

$$\min_{\tilde{\mathbf{p}}_n \in \Omega_n^{\tilde{P}}} \Pi_n(\tilde{\mathbf{p}}_n, \tilde{\mathbf{p}}_{-n}, \lambda_n, \mu) \tag{3.25}$$

The lagrangian (3.20) of each CR user in $G_{MPCF\text{-}S}$ contains two interference parts: one is the interference caused by the self power transmission; the other part is the interference caused from other CR users and PUs to it. Specifically, in the first part, each CR user need to adjust its transmit power that minimizes the desired transmit power under the target QoS and restricts its interference to all the PUs; in the second part, each CR user should transmit an adequate power to combat the Rayleigh-fading environments, and the interference from PUs and other CR users to guarantee its requested QoS.

3.4 Distributed Power-Control Algorithm

In this section, we design a distributed algorithm to find the Nash Equilibrium for G_{MPCF}. To obtain the global optimum of this problem, the gradient-descent method is applied intuitively. The gradient of $\Pi_n(\tilde{\mathrm{p}}_n, \tilde{\mathrm{p}}_{-n}, \lambda_n, \mu)$ with respect to $\tilde{p}_n^k$ is given by

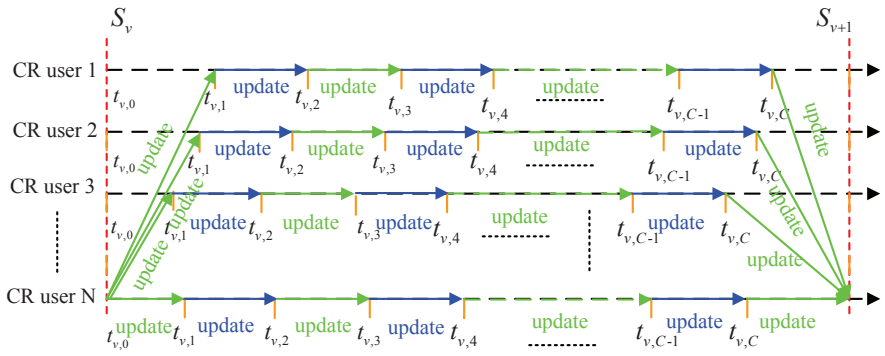

Figure 3.3. Asynchronous power control and dual-price updating process in ΔT.

$$\nabla\Pi_n(\tilde{\mathbf{p}}_n, \tilde{\mathbf{p}}_{-n}, \lambda_n, \mu) = \alpha_n e^{\tilde{p}_n^k} + \sum_{m=1}^{M} \mu_m^k \frac{b_{nm}^k e^{\tilde{p}_n^k}}{1 + b_{nm}^k e^{\tilde{p}_n^k}} - \lambda_n \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k e^{\tilde{p}_n^k}}$$
$$-\lambda_n \sum_{i \neq n}^{N} \frac{\bar{\gamma}_n^k g_{in}^k e^{\tilde{p}_i^k}}{g_{nn}^k e^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k e^{\tilde{p}_i^k}} - \lambda_n \sum_{j=1}^{M} \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k}{g_{nn}^k e^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k \bar{p}_j^k}.$$

In terms of the pricing interpretation of each term, it motivates the following asynchronous distributed algorithm, in which node pairs iteratively update their price announcements and

transmit-power decisions in an attempt to achieve the global optimum. For the layered structure of the power-updating algorithm, we denote ΔT as the long time updating interval for the dual price of both the PUs' and CR users' requested QoS. Let $\Delta t = \Delta T/C$ denote the short time-updating interval for the power-updating of the sub-game G_{MPCF-S}, where C is a constant and large enough such that for any given dual prices, $(\tilde{p}_1^*(\lambda,\mu),...,\tilde{p}_N^*(\lambda,\mu))$ of G_{MPCF-S} can be completed before the end of ΔT. The asynchronous power-control algorithm for sub-game G_{MPCF-S} requires no strict time synchronization among multiple CR users within the long time duration of $[S_v, S_{v+1}]$, which is more robust and suitable for the implementation in dynamic fading environments. In this case, each CR user can use other CR users' transmit power and channel information either before or after the current time slice within its power-updating process. As a simple illustration, the power-updating process of the CR user N is shown in Fig.3.3. The corresponding distributed power-control algorithm (**A1**) is given as follows.

Algorithm (A1): (***Asynchronous Distributed Power-Control Algorithm***)

Initialization:

At time $t_{0,0} = S_0 = 0$, each CR user $n \in \mathcal{N}$ chooses the initial $\tilde{p}_n(t_{0,0}) \in \Omega_n^P$ and set $\lambda_i(S_v) > 0$. Each PU $m \in \mathcal{M}$ initializes $\mu_m^k(S_0) > 0, \forall k \in \mathcal{K}$.

Iterative Process:

(I)**Sub-gradient Decent Updating for Power-Control Game** After receiving the dual prices $\lambda_i(S_v)$ and $(\mu_1(S_0),...,\mu_M(S_0))$, each CR user n will play the sub-game $G_{MPCF-S}(\lambda(S_v),\mu(S_v))$ during the v^{th} long updating interval. Specifically, during the v^{th} ΔT, each user $n \in N$ updates $\tilde{p}_n$ at the beginning of the c^{th} short updating time $t_{v,c}$ and initialize $\tilde{p}_n^k(t_{v,0}) = \tilde{p}_n^k(t_{v-1,c})$. The beginning of updating time $t_{v,c}$ for CR users do not need to be synchronized, but the end $t_{v,C}$ should be terminated before S_{v+1}.

$$\tilde{p}_n^k(t_{v,w}) = \tilde{p}_n^k(t_{v,w-1}) + \nu\nabla\Pi_n(\tilde{p}_n, \tilde{p}_{-n}, \lambda_n(S_v), \mu(S_v))$$

where $\lambda_n(t_{v,c})$ is the optimal dual price for CR user n's power budget constraint ($\sum_{k\in\mathcal{K}} p_n^k \leq P_n^{\max}$) at each $t_{v,c}$.

(II)**Sub-gradient Decent Updating for Dual Prices:**

Each CR user $n \in \mathcal{N}$ updates λ_n at the end of the v^{th} long updating time as follows:

$$\lambda_n(\mathrm{S}_{v+1}) = [\lambda_n(\mathrm{S}_v) + \varsigma_1 \sum_{k=1}^{\mathrm{K}} \{\sum_{i\neq n}^{N} \log(1 + \frac{\bar{\gamma}_n^k g_{in}^k \mathrm{e}^{\bar{p}_i^k(t_{v,C})}}{g_{nn}^k \mathrm{e}^{\bar{p}_n^k(t_{v,C})}}) + \sum_{j=1}^{M} \log(1 + \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k}{g_{nn}^k \mathrm{e}^{\bar{p}_n^k(t_{v,C})}}) + \log(1 - \xi_n^k) + \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k \mathrm{e}^{\bar{p}_n^k(t_{v,C})}}\}]^+$$

Each PU $m \in \mathcal{M}$ also updates μ_m at the end of the v^{th} long updating time as follows:

$$\mu_m^k(\mathrm{S}_{v+1}) = [\mu_m^k(\mathrm{S}_v) + \varsigma_2\{\sum_{n=1}^{N} (\log(1 + b_{nm}^k \mathrm{e}^{\tilde{p}_n^k(t_{v,C})}) - \log \frac{a_m^k}{1-\xi_m^k}\}]^+$$

(II) **Set v = v + 1 and Repeat Steps (a), (b), (c) until Convergence**

3.5 Convergence

In this section, the convergence of the proposed asynchronous algorithm is presented. According to [72], the totally asynchronous iteration allows CR users to perform power adjustments faster and execute more iterations than the synchronous iteration. It updates CR users' outdated information including the reception failures resulted from fading environments.

For example, define $\tau_n^j(t_{v,c})$ as the most recent time that $\tilde{\mathbf{p}}_n$ is known to CR user j at time $t_{v,c}$. Then the CR user j adjusts its transmit power at time $t_{v,c}$ using the power vector

$$\tilde{p}(\tau^j(t_{v,c})) := (\tilde{p}(\tau_1^j(t_{v,c})), ..., \tilde{p}_N(\tau_N^j(t_{v,c}))) \tag{3.26}$$

where $\tilde{p}_n(t_{v,c})$ is the transmitted power of CR user n at time $t_{v,c}$. Let $T^n(t_{v,c})$ be the time sets at which $\tilde{p}_n(t_{v,c})$ is updated. Denote the update algorithm as $\mathcal{T}_n(\cdot)$. Given the sets $T^1, ..., T^n, ..., T^N$, the totally asynchronous standard algorithm model can be defined as

$$\tilde{p}_n(t_{v,c+1}) = \begin{cases} \mathcal{T}_n(\tilde{p}(\tau^n(t_{v,c}))), & t \in T^n; \\ \tilde{p}_n(t_{v,c}), & otherwise. \end{cases} \tag{3.27}$$

The sets T^j are infinite at a given long-time interval v, and given any $t_{v,0}$, there exists $t_{v,1}$ such that $\tau_n^j(t_{v,c} \geq t_{v,0})$, for all $t_{v,c} \geq t_{v,1}$.

The convergence of the proposed power-control algorithm can be proved by the *Asynchronous Convergence Theorem* in [72] as stated in the following theorem 2.

Theorem 2 If there is a sequence of nonempty sets $\{\mathcal{P}(i)\}$ with $\{\mathcal{P}(i+1)\} \subset \{\mathcal{P}(n)\}$ and $\mathcal{T}(\tilde{\mathbf{p}}) = (\mathcal{T}_n(\tilde{\mathbf{p}})), n \in N$ for all i satisfying the following two conditions:

(1) **Synchronous Convergence Condition**: for all i and $\tilde{\mathbf{p}} \in \mathcal{P}(n), \mathcal{T}(\tilde{\mathbf{p}}) \in \mathcal{P}(i+1)$. If q^i is a sequence such that $q^i \in \mathcal{P}(i)$, $\tilde{\mathbf{p}}$ is a fixed point of $\mathcal{T}(\tilde{\mathbf{p}})$.

(2) **Box Condition**: For every i, there exist sets $\mathcal{P}_n(i) \in \mathcal{P}$, such that $\mathcal{P} = \mathcal{P}_1(i) \times \mathcal{P}_2(i) \times ... \times \mathcal{P}_N(i)$.

For the proposed algorithm, $\mathcal{T}(\tilde{\mathbf{p}})$ is depicted as (3.26). Since the proposed game (Theorem 1) has convexity, the fixed point $\tilde{\mathbf{p}}^\star$ of $\mathcal{T}(\tilde{\mathbf{p}})$, i.e. $\tilde{\mathbf{p}}^\star = \mathcal{T}(\tilde{\mathbf{p}}^\star)$ is also the global optimum of the game. The challenges in applying Theorem 2 are to identify a suitable sequence of sets $\{\mathcal{P}(i)\}$. The concept of contraction mapping [72] plays a key role in the proof of the following theorem.

Theorem 3 *[Convergence]* For the sufficient small ν, the proposed distributed algorithm converges to the global optimum $\tilde{p}^\star$ when the power-control algorithm operated asynchronously for the sub-game $\mathrm{G}_{MPCF\text{-}S}$.

Proof To prove this theorem, it is enough to prove the proposed algorithm $\mathcal{T}(\tilde{\mathbf{p}})$ is a special type of pseudo-contraction mapping [72], that is

$$\left\|\mathcal{T}(\tilde{\mathbf{p}}) - \tilde{\mathbf{p}}^\star\right\| \leq \left\|\tilde{\mathbf{p}} - \tilde{\mathbf{p}}^\star\right\|. \tag{3.28}$$

where $\|\cdot\|$ is certain norm, and $\rho \in [0, 1)$ is a constant. By choosing the sets of sequence $\{\mathcal{P}(i)\}$ in theorem 2 as

$$\mathcal{P}(i) = \{\tilde{\mathbf{p}} | \left\|\tilde{\mathbf{p}} - \tilde{\mathbf{p}}^\star\right\| \leq \rho^n \left\|\tilde{\mathbf{p}}(0) - \tilde{\mathbf{p}}^\star\right\|\}. \tag{3.29}$$

It follows the condition (1) and (2) in theorem 2. Therefore, all we need is to prove $\mathcal{T}(\tilde{\mathbf{p}})$ is a special type of pseudo-contraction mapping. For the clear description of this point, we first prove the conclusion for the single-channel case, $K = 1$, where the channel subscripts k are omitted, and the conclusion is extended to $K > 1$ later.

Let $f(\tilde{\mathbf{p}}) = \nabla\Pi_n(\tilde{\mathbf{p}}_n, \tilde{\mathbf{p}}_{-n}, \lambda_n, \mu)$, and $\nabla_i f_n(\tilde{\mathbf{p}})$ denotes the partial derivatives of components f_n with respect to the component of $\tilde{p}_i$. Accordingly, we have

$$\begin{aligned}|1-\nu\nabla_n f_n(\tilde{p})| = |1-\nu\{\alpha_n e^{\tilde{p}_n^k} + \sum_{m=1}^{M}\mu_m^k \frac{b_{nm}^k e^{\tilde{p}_n^k}}{(1+b_{nm}^k e^{\tilde{p}_n^k})^2} + \lambda_n \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k e^{\tilde{p}_n^k}} \\ + \lambda_n \sum_{i\neq n}^{N} \frac{\bar{\gamma}_n^k g_{in}^k e^{\tilde{p}_i^k} * g_{nn}^k e^{\tilde{p}_n^k}}{(g_{nn}^k e^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k e^{\tilde{p}_i^k})^2} + \lambda_n \sum_{j=1}^{M} \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k * g_{nn}^k e^{\tilde{p}_n^k}}{(g_{nn}^k e^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k)^2}\}.\end{aligned} \tag{3.30}$$

Therefore, $1-\nu\nabla_n f_n(\tilde{p}) > 0$ for ν is small enough. Similarly, we have

$$\sum_{i\neq n} |\nu\nabla_i f_n(\tilde{p})| = \nu|\lambda_n \sum_{i\neq n} \frac{\bar{\gamma}_n^k g_{in}^k e^{\tilde{p}_i^k} * g_{nn}^k e^{\tilde{p}_n^k}}{(g_{nn}^k e^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{in}^k e^{\tilde{p}_i^k})^2}|. \tag{3.31}$$

So, we have

$$|1-\nu\nabla_n f_n(\tilde{p})| + \sum_{i\neq n} |\nu\nabla_i f_n(\tilde{p})| < \varepsilon \tag{3.32}$$

where

$$\begin{aligned}\varepsilon = 1-\nu\{\alpha_n e^{\tilde{p}_n^k} + \sum_{m=1}^{M}\mu_m^k \frac{b_{nm}^k e^{\tilde{p}_n^k}}{(1+b_{nm}^k e^{\tilde{p}_n^k})^2} + \lambda_n \frac{\bar{\gamma}_n^k \sigma_k^2}{g_{nn}^k e^{\tilde{p}_n^k}} \\ + \lambda_n \sum_{j=1}^{M} \frac{\bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k * g_{nn}^k e^{\tilde{p}_n^k}}{(g_{nn}^k e^{\tilde{p}_n^k} + \bar{\gamma}_n^k g_{jn}^k \bar{p}_j^k)^2}\}\end{aligned} \tag{3.33}$$

Since all the terms in $\{\cdot\}$ are positive, there must exist a ν small enough such that $\varepsilon < 1$, i.e. the matrix of $\nabla f(\tilde{p})$ is diagonal dominant. From [72], we know the proposed algorithm is a pseudo-contraction mapping when $K = 1$.

For $K > 1$, the scalar $\tilde{p}_n$ is replaced by vector $\tilde{\mathbf{p}}_n = (\tilde{p}_n^k, k \in K)$, so the $\nabla_n f_n(\tilde{p})$ denotes a matrix of dimension $K * K$. Now, the $|\cdot|$ is replaced by the *spectrum norm*. Similarly, the contraction mapping can be proved by the diagonal-dominance condition. The detailed proof for multiple channels cases is omitted here for brevity.

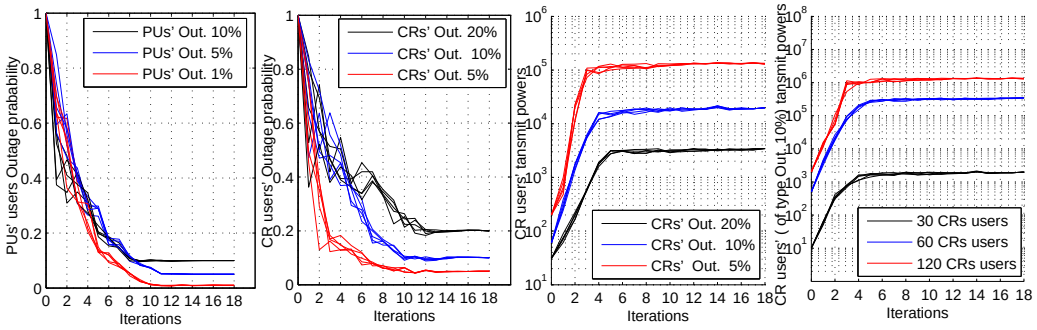

Figure 3.4. Convergence process of outage probability and transmit powers for CR users and/or PUs.

3.6 Simulation Results

We consider a network where PUs and CR users are within a $1km$-radius circle in the simulation. The location of both PUs and CR users follows the uniform distribution. We investigate the performance of the power-control algorithm of CR users, and channel outages of both PUs and CR users over six shared spectrum bands. The weighting factor $\alpha_n = 1, \forall n \in \mathcal{N}$. The channel gain of i^{th} user is determined by the Rayleigh fast-fading and a log-normal shadowing path-loss model, given by $g_i = (8/d_i)^4 \cdot Y_\sigma^{-1} \cdot f_i$, where d_i is the distance-dependent with loss exponent 4, $\text{Log}(Y_\sigma)$ is a zero-mean Gaussian random variable with a standard deviation of $\sigma = 8$, and f_i is a random variable with Rayleigh distribution. The variable f_i models the fast-fading channel. We generate the random variable f_i at each time (or iteration) step, and Y_σ very 30 time steps. Furthermore, an AWGN noise power is chosen to be $\sigma^2 = 10^{-13}$, corresponding to an approximate $1MHZ$ bandwidth.

We set the default simulation parameters as follows: there are three types of CR users and PUs, each having the typical outage-probability and SINR-threshold pairs at receiver side $(5\%, 12dB)$, $(10\%, 10dB)$, $(20\%, 8dB)$ and $(1\%, 10dB)$, $(5\%, 8dB)$, $(10\%, 6dB)$, respectively. Each PU's transmit power is assumed to be $80dBm$ (i.e. the typical transmission power of primary base (FM radio) stations). Typically, we consider 30 CR users and 30 PUs, and both of them have 10 users of each type. Our proposed asynchronous distributed power-control algorithm 1 is implemented in Matlab, and it iterates to declare the convergence until successive iterations yield an average per-user power within 1%.

Fig. 3.4. (1) (2) show the convergence process of the outage probability for CR users and PUs, respectively. When comparing the convergence process of CR users and PUs, the PUs' outage probability converges smoother than the CR users. In Fig.3.4. (3), (4), the convergence process of the transmit power is shown for the CR users with different *type* and *number* of users. We consider 100 independent simulations for each type of users, and obtain the averaged values of transmit power for each type of users. Fig.3.4. (3) represents that the lower the channel-outage constraints, the higher the transmit power

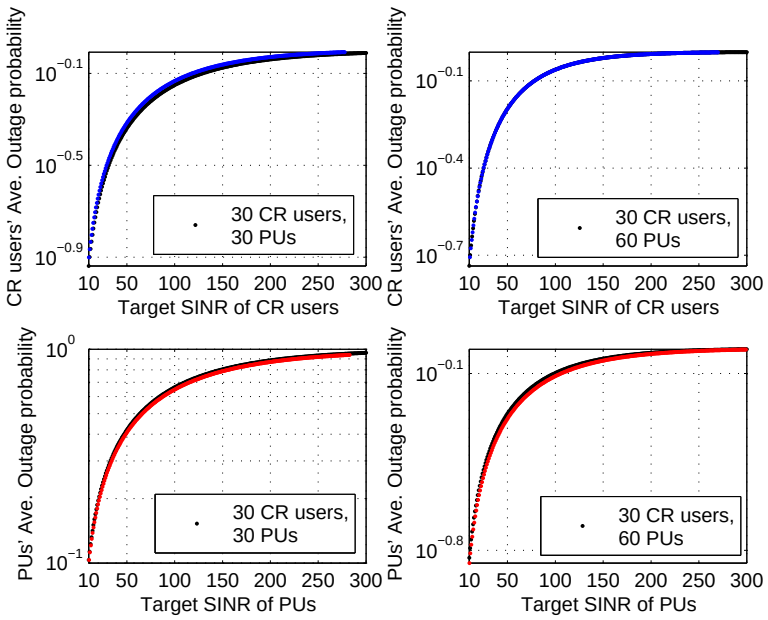

Figure 3.5. SINR *vs* outage probability for CR users and PUs.

required for each type of CR users. In addition, we can observe the fact that increasing number of users results in the transmit power increment in Fig.3.4. (4). All the results confirm our theoretical predictions. Moreover, we could see that the transmit power and outage probability of each CR and/or PU do not monotonically converge to the equilibrium points. The rationale on this point is, in the present of Rayleigh fading, the channel gain is varying with time, which affects the choice of power strategies in the power-control game.

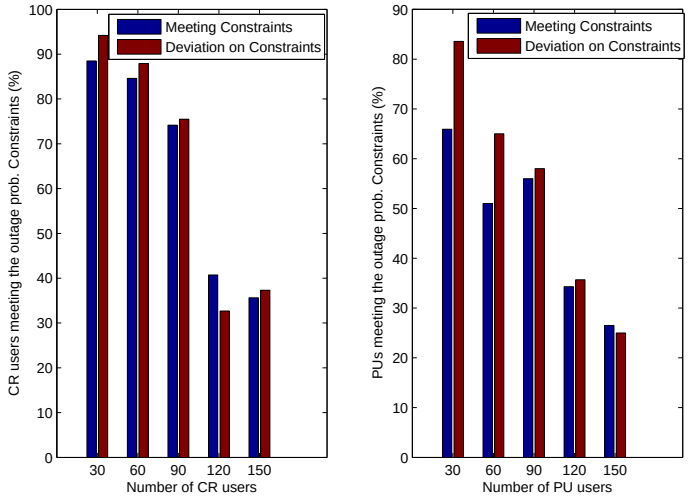

Figure 3.6. Percentage of meeting outage-probability constraints.

Fig.3.5 shows the outage-probability constraints with respect to the different target SINRs for both CR users and PUs. The target SINR is changing from 12 ($10dB$) to 300 ($24dB$). As the target SINR increases, the outage probability always increases at the same time. In this figure, both theoretical results and the simulation results of the average outage probability are shown for CR users. Further quantification can be performed to the average deviation of specified outage probability and the simulated outage probability in Fig.3.6. Furthermore, we can observe the enhancement of the constraints' deviation, when the size of the users (PUs or CR users) is scaled upwards.

CHAPTER 4

LEARNING THROUGH REINFORCEMENT FOR REPEATED POWER-CONTROL GAME IN COGNITIVE RADIO NETWORKS

4.1 Repeated Constrained Power-Control Game

4.1.1 System Model as learning Automaton

The system model is illustrated in Fig. 1, where there are N CR transceiver pairs playing the repeated power-control game with PUs over time. There are M PUs in the network region. Each PU's behavior might be changing with time, where its activity is distributed according to some probability that is not known to CR users. The interference CSI condition among CR users and the CSI from each CR user to each PU are unknown, and they might be changing with time. We consider a the synchronous slotted time structured spectrum access for PUs and CR users as in [56] [55] during the long time repeated game. At each time slot t, each CR transmitter forwards its data with transmit power strategies from $\mathsf{P}_k^t \in \mathbf{P}_k \triangleq \{\mathsf{P}_k(1), \mathsf{P}_k(2), ..., \mathsf{P}_k(N_k)\}$, which is the action set of the learning automaton of user k; and N_k is the number of total actions of the CR user k. The transmit power space for the CRN is denoted as $\mathbf{P} = \prod_{k=1}^{N} \mathbf{P}_k$. CR users can utilize its feedback channel (Rx-to-Tx) to estimate its own CSI. Let us denote $s_k^t = (b_l^t, G_{kk}^t)$ as the CSI of the CR user k at time t, where $b_l^t = 0$ ($b_l^t = 1$) indicates the inactive (active) of PU l's transmission at time t, and G_{kk}^t denotes the channel gain of the CR user k. Since we assume the blind CSI from other CR users and PUs, the expected payoff u_k^t and the interference function (from the CR user k to the PU l) $\eta_{k,l}^t(b_l^t)$ for each CR user is unknown before the game.

For the PUs' communications, after every primary data transmission, PU-Rxs will feedback an interference term that includes the background noise and the collection of the sum-interference made from CR users. This information is used by PUs to monitor the interference level of the environment for the notification of the violation of interference power constraint messages[1] to CR networks and to conduct their own transmit power control. CR

[1]This is intrinsically the same as in [56] [55]. They assume the PUs measures all the sum-interference

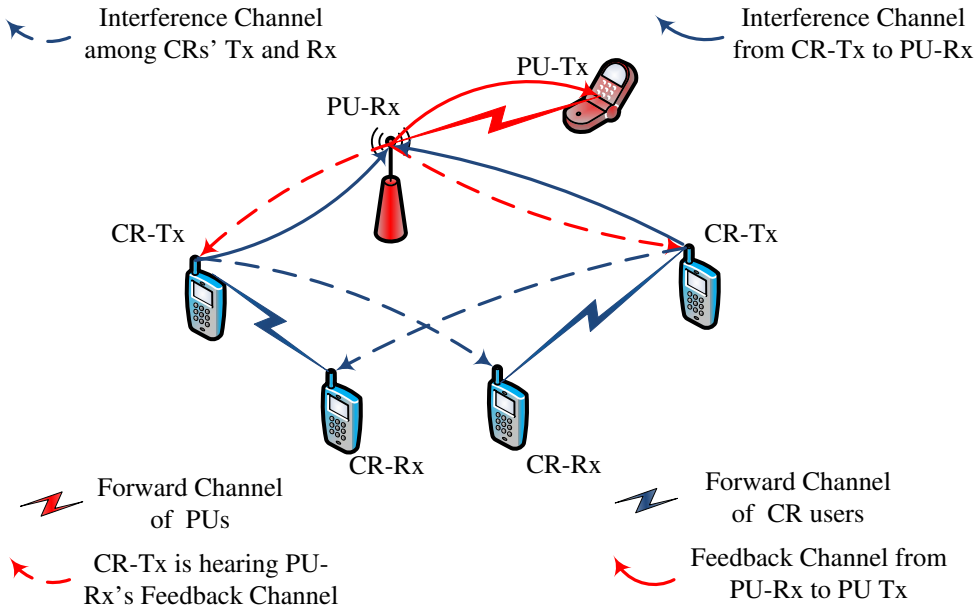

Figure 4.1. CR users explore PUs-RX's feedback information; they don't have channel and power strategy information from other CR users and to PUs.

users are assumed to have the capability to explore this feedback information from the control links of PUs as in [29] [56] [55]. The exploration process of this feedback information can have errors and contain not updated values some time[2]. As for the wireless environments, the wireless channel can be modeled as a finite-state Markov channel (FSMC) [73], where fading can be approximated as a discrete-time Markov process with time discretized to a given time interval [74]. Thus, the behavior of each player (or user) in the wireless fading environment can be modeled by a *stochastic variable learning automation* [75] that consists a Markov chain of finite states.

Let (Ω, H, P) be a probability space, where Ω is the sample space; H is a minimal σ-algebra on subsets of Ω; and P is a probability measure on (Ω, H). The symbol ω denotes an event in the probability space Ω. The stochastic automaton operating in a dynamic wireless

from all CR users, and the PU-Rx calculate and feedback the channel outage ACK/NACK to the PU-Tx in fading environments. Our assumption is of the same functionality as in [56] [55], but ours is much more suitable for general wireless environments that includes fading and no-fading conditions (See Definition 1 in Section I). The feedback background noise (sum-interference) information is a value of several bits that used for the power control and interference management in the PU-Tx side.

[2]In the following analysis, this means the probability of error exploration of the interference level approaches to 0 almost surely, mathematically.

environment is a adaptive discrete machine described by the tuple $\{\Xi, \mathbf{P}_k, \{\mathsf{P}_k^t\}, \{s_k^t\}, \{u_k^t\}, \{\eta_k^t(b_k^t)\}, \{p_k^t\}\}$, where Ξ is the automation input bounded set; $\mathbf{P}_k, \{\mathsf{P}_k^t\}, \{s_k^t\}, \{u_k^t\}, \{\eta_{k,l}^t(b_l^t)\}$ are, the action set of transmit power automaton (as we defined above), the sequences of automaton outputs (actions) of transmit power, input channel states, input utilities/payoffs and input interference functions, respectively. Let $p_k^t \stackrel{\Delta}{=} [p_k^t(1), p_k^t(2), ..., p_k^t(N_k)]^T$ denote the mixed strategy of transmit power of user k at time t, where N_k is the number of transmit power strategies of user k. The mixed strategy p_k^t is the conditional probability distribution at time t

$$p_k^t(i) = \Pr\{\omega \in \Omega : \mathsf{P}_k^t = \mathsf{P}_k^t(i)/H_{t-1}\}, \sum_{i=1}^{N_k} p_k^t(i) = 1.$$

where $\mathrm{H}^t = \sigma(\mathsf{P}_k^1, s_k^1, u_k^1, \eta_{k,l}^1(b_l^1), p_k^1; ...; \mathsf{P}_k^t, s_k^t, u_k^t, \eta_{k,l}^t(b_l^t), p_k^t;$
$b_k^t = 0, 1, k = 1, .., N, l = 1, .., M)$ is a σ-algebra that generated by all the histories event over the action sets of CR users. Let $\mathrm{T}^t = \mathrm{T}_k^t$ represent a reinforcement learning (updating) scheme of each CR user which changes the probability vector p_k^t to p_k^{t+1}, that is

$$\mathrm{p}_k^{t+1} = \mathrm{p}_k^t + \gamma_k^t \mathrm{T}^{\mathrm{t}}, \tag{4.1}$$

where $\mathrm{T}^t = \mathrm{T}^t(p_k^t; \{s_k^t\}^{t=1,...,t}, \{u_k^t\}^{t=1,...,t}, \{\eta_k^t(b_k^t)\}^{t=1,...,t}$
$, \{\mathsf{P}_k^t\}^{t=1,...,t})$, and γ_k^t is a scaler correction factor. The vector $\mathrm{T}_k^t(.) = [\mathrm{T}_{k,1}^t(.), ..., \mathrm{T}_{k,N_k}^t(.)]^T$ satisfies the following conditions that guarantees the valid probability distribution:

$$\sum_{i=1}^{N_k} \mathrm{T}_{k,i}^t(.) = 0, \quad p_k^t(i) + \gamma_k^t \mathrm{T}_{k,i}^t(.) \in [0, 1]. \tag{4.2}$$

for any t and $i = 1, ..., N_k$. The above is the kernel of the the *learning automation* (LA).

4.1.2 Repeated Game Behavior

Non-cooperative power control is inherently a repeated process. Each user collects locally the information from the network and then makes the decision of transmit power levels locally. When new information (perhaps dependent on actions in the prior periods) is available, the decision process repeated itself. The learning-based repeated power-control game is played in the following way. Consider a noncooperative game in which each player or

CR user has only known its own payoff (or utility) realization u_k^t, and $\eta_l^t(b_l^t)$ the realization of interference function (interference term) explored from PUs after the game. At each stage (or time) t, each CR user will simultaneously and independently choose an action according to the probability distribution (mixed strategy) p_k^t, for instance, the transmit power $\mathsf{P}_k(i_k), 1 \le i_k \le N_k, 1 \le k \le N$. Let the achievable data transmission rate r_k^t of CR user k applied to the instant utility:

$$\begin{aligned} r_k^t(\mathbf{P}^t) &= W\log_2(1 + K\mathsf{SINR}(\mathbf{P}^{\mathrm{t}})) \\ &= W\log_2(1 + \frac{K\mathsf{P}_k^t G_{kk}(s_k^t)}{\sum_{j\neq \mathrm{k}}^{j\in N} \mathsf{P}_j^t G_{kj}(s_j^t) + \mathrm{n}_k^t}) \\ &= W\log_2(1 + \frac{K\mathsf{P}_k^t G_{kk}(s_k^t)}{\sum_{j\neq \mathrm{k}}^{j\in N} \mathsf{P}_j^t G_{kj}(s_j^t) + \sigma_k^2 + \sum_{l=1}^{M} \delta_{k,l} b_l^t}) \end{aligned}, \tag{4.3}$$

where $\mathbf{P}^t = (\mathsf{P}_1^t, \mathsf{P}_2^t, .., \mathsf{P}_N^t) \in \mathbf{P}$. The symbol W denotes the bandwidth of the primary spectrum. K is the SINR gap between a practical data transmission rate and the Shannon capacity limit. The function $G_{kj}(s_j^t)$ represents the channel gain from CR user k to CR user j in channel state s_j^t. The factors $G_{kj}, 1 \le j \le N$, incorporate propagation loss (slow-varying and fast-varying parts), spreading gain, and other normalization constants. Notice that G_{kk} is the path gain on link k. Assuming that not too many close-by nodes transmit at the same time, with reasonable spreading gain, G_{kk} is much larger than other G_{kj}, $(G_{kk} \gg G_{kj})$ $j \neq k$. This assumption is the same as in [76], and we denote it as the *interference non-dominance* condition. The notation n_k^t represents the instant noise of CR user k at time t, in which σ_k^2 is the noise power under the assumption of additive white Gaussian distribution, $\delta_{k,l}$ is the interference from PU l's transmission imposed on CR user k, and b_l^t is the indicator function of the existence of primary transmission.

Since each CR user only know its own transmit power strategy, the mixed instant utility of CR user k is given as follows

$$\mathrm{u}_k^t(p_k^t) = \sum\nolimits_{i_k=1}^{N_k} \mathrm{r}_k^t(\mathsf{P}_k^t(i_k)) \cdot p_k^t(i_k). \tag{4.4}$$

Similarly, the instant and the mixed instant interference function of CR user k imposed on

PU l is given by

$$\eta_{k,l}^{t}(\mathsf{P}_{k}^{t}(i_k)) = \mathsf{P}_{k}^{t} G_{kl}(s_{k}^{t}). \tag{4.5}$$

$$\eta_{k,l}^{t}(p_{k}^{t}) = \sum\nolimits_{i_k=1}^{N_k} b_{l}^{t} \cdot \eta_{k,l}^{t}(\mathsf{P}_{k}^{t}(i_k)) \cdot p_{k}^{t}(i_k). \tag{4.6}$$

Then, for each CR user, it changes its mixed strategy according to the accepted reinforcement $\mathrm{p}_{k}^{t} \xrightarrow{T_k^t} \mathrm{p}_{k}^{t+1}$ trying for large t to maximize its individual averaged utility $\mathrm{liminf}_{t\to\infty}\, \mathrm{u}_{k}^{t}$ where

$$\Phi_{k}^{t} = \frac{1}{t} \sum\nolimits_{\tau=1}^{\tau=t} \mathrm{u}_{k}^{\tau}(p_{k}^{\tau}), \tag{4.7}$$

and the sum-interference of all CR users measured by the PU l at time t is

$$\eta_{l}^{\tau}(\mathbf{p}^{\tau}) = \sum\nolimits_{k=1}^{k=N} \eta_{k,l}^{\tau}(p_{k}^{\tau}). \tag{4.8}$$

Its time average should conform to the interference power constraint C_l. If the *interference-power-constraint-violation* information is reported in its feedback control channel at time t, this information (eavesdropped by CR users) will reinforce the CR users to change their power control strategies in the long run to conform to C_l:

$$\Psi_{l}^{t} = \frac{1}{t} \sum\nolimits_{\tau=1}^{\tau=t} \eta_{l}^{\tau}(\mathbf{p}^{\tau}), \mathrm{liminf}_{t\to\infty}\, \Psi_{l}^{t} \leq C_l. \tag{4.9}$$

Note that in the above learning game, each individual CR user k only known its own utility and the interference power level that is eavesdropped from PUs' feedback control link after the play in each step. It chooses the transmit power strategy $\mathsf{P}_{k}^{t}(i_k)$ with probability $p_{k}^{t}(i_k)$ according to the reinforcement learning scheme described in the previous subsection.

As we know, in most researches in wireless communications, the wireless channels or environments are assumed to be stationary. Hence this assumption on a sequence of real functions (e.g., utility and interference function in this work) is just one special case that needed in order for the results in the stochastic analysis to be true. In practice, the wireless environment might experience with non-stationary observation of noise and interference. Let us consider the non-stationary environment which is characterized by the following two properties:

Assumption 1: The conditional expectations of u_k^t and η_l^t of the environment responses exist, i.e.,

$$\begin{aligned}
&\mathrm{E}\{\mathrm{u}_k^t|\mathsf{P}_k^t{=}\mathsf{P}_k(\mathrm{i}_k) \wedge \mathrm{H}_{\mathrm{t}-1}, \forall \mathrm{k} \in \mathrm{N}\} = r_k^{i_1,\ldots,i_N}(t) \\
&\quad = W\log_2(1 + K\mathsf{SINR}(\mathsf{P}^t(i_1), \ldots, \mathsf{P}^t(i_k), \ldots, \mathsf{P}^t(i_N))) \\
&\mathrm{E}\{\eta_l^t|\mathsf{P}_k^t{=}\mathsf{P}_k(\mathrm{i}_k) \wedge \mathrm{H}_{\mathrm{t}-1}, \forall \mathrm{k} \in \mathrm{N}\} = \eta_l^{i_1,\ldots,i_N}(t) \\
&\quad = b_l^t \cdot \left(\sum_{\mathrm{k}=1}^{N} \sum_{i_k=1}^{N_k} \mathsf{P}_k^t(\mathrm{i}_k) G_{kl}(s_l^t) \cdot p_k^t(i_k) + \sigma_l^2 \right).
\end{aligned} \tag{4.10}$$

and their arithmetic averages tend to a finite limit, with probability one,i.e.,

$$\begin{aligned}
&\tfrac{1}{t} \sum_{\tau=1}^{\tau=t} r_k^{i_1,\ldots,i_N}(\tau) \underset{t\to\infty}{\overset{a.s.}{\to}} r_k^{i_1,\ldots,i_N} \\
&\tfrac{1}{t} \sum_{\tau=1}^{\tau=t} \eta_l^{i_1,\ldots,i_N}(\tau) \underset{t\to\infty}{\overset{a.s.}{\to}} \eta_l^{i_1,\ldots,i_N}.
\end{aligned} \tag{4.11}$$

The notation $r_k^{i_1,\ldots,i_N}(t)$ represents the expected transmission rate of the CR user k given a specific transmit power strategy of $i_1, \ldots, i_N$ for all the N CR users. Similarly, we have the expected interference level $\eta_l^{i_1,\ldots,i_N}(t)$ measured by the PU l. The first part of the assumption A1 (4.10) states some kind of restriction to the properties of the observation noise. Take the η_l^t in (4.10) as an example, the condition will be satisfied if at every time t the noise σ_l are bounded in second moments, i.e, $\mathrm{E}\{\sigma_1^2\} < \infty$. This property holds, for example, for Gaussian noises (it is assumed in (4.3)) and is not true for noise having Cauchy distribution. The second property of assumption A1 (4.11) represents some kind of stationary in average. It will be satisfied for the utility, interference function and noise, which are stationary in average.

Assumption 2: For any realized action of the transmit power P_k^t at current time t in the repeated power-control game, the conditional variance of the utility u_k^t and the interference functions η_l^t of the CR user k are uniformly bounded with probability one, i.e.,

$$\begin{aligned}
&E\left\{ \left(u_k^t - r_k^{i_1,\ldots,i_N}(t) \right)^2 / H_{t-1} \wedge \mathsf{P}_k^t{=}\mathsf{P}_k(\mathrm{i}_k) \right\} = \sigma_{t,k}^+, \\
&\qquad \max_k \sup_t \sigma_{t,k}^+ = \sigma_{u,k}^+ < \infty, \forall k = 1, \ldots, N \\
&E\left\{ \left(\eta_l^t - \eta_l^{i_1,\ldots,i_N}(t) \right)^2 / H_{t-1} \wedge \mathsf{P}_k^t{=}\mathsf{P}_k(\mathrm{i}_k) \right\} = \sigma_{t,l}^+, \\
&\qquad \max_k \sup_t \sigma_{t,l}^+ = \sigma_{\eta,l}^+ < \infty, \forall k = 1, \ldots, N
\end{aligned} \tag{4.12}$$

The notation $\sigma^{+}_{u,k}$ and $\sigma^{+}_{\eta,l}$ represent the upper bounds of the transmission rate of the CR user k and interference function measured by the PU l. Since the number of transmit power strategies for each CR user is finite and the channel gains of wireless environment (including any fading models) are bounded almost surely [74], the upper bounds $\sigma^{+}_{u,k}$ and $\sigma^{+}_{\eta,l}$ must exist.

Definition 1: A random environment, satisfying conditions (4.10)~(4.12) is said "***asymptotically stationary in the average sense***".

For each CR user, the average utility and interference function from the collection of 2 tensors $U_k = [r_k^{i_1,\ldots,i_N}(t)], (1 \le k \le N, 1 \le i_k \le N_k)$ and $Q_l = [\eta_l^{i_1,\ldots,i_N}(t)], (1 \le l \le M, 1 \le i_k \le N_k)$ which are *priori unknown*. Thus, we can define the expected utility and expected constraints concerning with all PUs' and/or CR users' CSI and transmit power strategy information as follows.

Definition 2*:* At time t, for the CR user k in the considered game, and for any H_{t-1}-measurable conditional probability distribution $(\mathrm{p}_1^t, ..., \mathrm{p}_N^t)$, the *expected utility* is defined as

$$\begin{aligned} U_k(p_1^t, ..., p_N^t) &= \mathrm{E}[r_k^{i_1,\ldots,i_N}(t)] \\ &= \sum_{i_1=1}^{N_1} \ldots \sum_{i_N=1}^{N_N} r_k^{i_1,\ldots,i_N}(t) \prod_{s=1}^{N} p_s^t(i_s) \\ &= W\log_2(1 + \frac{K\mathsf{P}_k^t \cdot p_k^t G_{kk}(s_k^t)}{\sum_{j\neq k}^{j\in N} \mathsf{P}_j^t \cdot p_j^t G_{kj}(s_j^t) + \mathrm{n}_k^t}) \\ &= W\log_2(1 + \frac{K\sum_{i_k=1}^{N_k} \mathsf{P}_k^t(i_k) p_k^t(i_k) G_{kk}(s_k^t)}{\sum_{j\neq k}^{j\in N} \sum_{i_j=1}^{N_j} \mathsf{P}_j^t(i_j) p_j^t(i_j) G_{kj}(s_j^t) + \mathrm{n}_k^t}) \end{aligned},$$

where $\mathsf{P}_j^t \cdot p_j^t = \sum_{i_j=1}^{N_j} \mathsf{P}_j^t(i_j) p_j^t(i_j), 1 \le j \le N$ is the expected transmit power of the mixed strategy of the CR user j at time t, and the corresponding *expected interference function* is defined as

$$\begin{aligned} Q_l(p_1^t, ..., p_N^t) &= \mathrm{E}[\eta_l^{i_1,\ldots,i_N}(t)] \\ &= \sum_{i_1=1}^{N_1} \ldots \sum_{i_N=1}^{N_N} \eta_l^{i_1,\ldots,i_N}(t) \prod_{s=1}^{N} p_s^t(i_s) \\ &= b_l^t \left(\sum_{\mathrm{k}=1}^{N} \mathsf{P}_k^t \cdot p_k^t G_{kl}(s_l^t) + \sigma_l^2 \right) \\ &= b_l^t \left(\sum_{\mathrm{k}=1}^{N} \sum_{i_k=1}^{N_k} \mathsf{P}_k^t(i_k) p_k^t(i_k) G_{kl}(s_l^t) + \sigma_l^2 \right) \end{aligned}.$$

We also have the *averaged expected utility* and *interference functions* over time t as follows:

$$\begin{aligned} U_k^t &= \frac{1}{t}\sum_{\tau=1}^{\tau=t} U_k{}^{\tau}(p_1^{\tau}, ..., p_N^{\tau}), \\ Q_l^t &= \frac{1}{t}\sum_{\tau=1}^{\tau=t} Q_l^{\tau}(p_1^{\tau}, ..., p_N^{\tau}). \end{aligned} \tag{4.13}$$

The following lemma indicates the asymptotic equivalence between (Φ_k^t, Ψ_l^t) and (U_k^t, Q_l^t).

Lemma 1: Under the assumption 1, for any conditional distribution $(\mathrm{p}_1^t, ..., \mathrm{p}_N^t)$, it has

$$\Phi_k^t = U_k^t + o_\omega(t^{-1/2}), \Psi_l^t = Q_l^t + o_\omega(t^{-1/2}).$$

Proof The proof is obtained in the view of Borel-Cantelli lemma [77] and the strong law of large numbers for dependent sequence [78]. $o_\omega(t^{-1/2})$ is a random sequence tending almost surely to zero, and more quickly than $t^{-1/2}$.

4.2 Problem Formulation

4.2.1 Mixed Strategies and Nash Equilibrium

The mixed power control strategy of the CR user k is any sequence of vectors $\mathrm{W}_k = \{p_k^t | \sum_{i=1}^{N_k} p_k^t(i) = 1, 1 \le k \le N\}$ with H_{t-1} measurable components belonging to the simplex $S_{\varepsilon=0}^{N_k}$, i.e.,

$$p_k^t \in S_{\varepsilon=0}^{N_k} = \{p_k^t \in R^{N_k} : p_k^t(i) \ge \varepsilon \ge 0, \sum_{i=1}^{N_k} p_k^t(i) = 1\}. \tag{4.14}$$

Definition 3: The power control strategies $W_1^*, ..., W_N^*$ in the noncooperative game are defined as Nash Equilibrium if:

1) The power control strategies are *admissible*, where the interference function Q_l is less than the PU l's interference power constraint C_l, i.e., for any $1 \le l \le M$

$$\max_{p_s^1, s\in N} \limsup_{t\to\infty} Q_l^t = \tilde{Q}(W_1^*, ..., W_N^*) \le C_l, \tag{4.15}$$

where C_l are *priori* know values[3];

2) for $\forall \mathrm{k} \in N$ and any admissible strategy W_k

$$U(W_1^*, ..., W_N^*) = \min_{p_s^1, s\in N} \liminf_{t\to\infty} U_k^t \ge U(W_1^*, ...W_k..., W_N^*), \tag{4.16}$$

[3]Or can also be assumed unknown, in this case, when the CR users violate this constraint, PUs will send the interference-constraints-violation message, and CR users can explore the interference level.

where the maximization of (4.15) and the minimization of (4.16) are guaranteed over all the initial probability distributions $p_s^1, s \in N$.

Definition 4: The point $(p_1^*, ..., p_N^*)$ is said to be a Nash Equilibrium point of the given N-user CR power-control game within the class of mixed strategies W_k, if for each $1 \leq k \leq N$

$$\begin{aligned} &U_k(p_1^*, ..., p_N^*) = \max_{p_k \in R^k} U_k(p_1^*, ..., p_{k-1}^*, p_k, p_{k+1}^*, ..., p_N^*) \\ &s.t.\ R^k = \bigcap_{l=1}^{M} \{p_k : Q_l(p_1^*, ..., p_{k-1}^*, p_k, p_{k+1}^*, ..., p_N^*) \leq C_l\}. \end{aligned} \tag{4.17}$$

At the Nash Equilibrium point, no player can increase his payoff by a unilateral change in its strategy.

Definition 5: A mixed strategy is said to be *stationary*, where $\mathrm{T}_k = \{p_k | \sum_{i=1}^{N_k} p_k(i) = 1, 1 \leq k \leq N\}$ is not changing with time after the learning of the repeated game process. The set of distributions $(p_1^*, ..., p_{\mathrm{N}}^*)$ is called the *stationary* mixed strategy of Nash Equilibrium.

Remark 1: The set of all equilibrium strategies W_k^*, where $1 \leq k \leq N$, contains the subsets of *admissible stationary strategy* $\{p_k^*\}$ for any $1 \leq k \leq N, 1 \leq l \leq M$ and for any $p_k \in S_{\varepsilon=0}^{N_k}$ such that,

$$\begin{aligned} &U_k(p_1^*, ..., p_N^*) \geq U_k(p_1^*, ..., p_{k-1}^*, p_k, p_{k+1}^*, ..., p_N^*). \\ &Q_l(p_1^*, ..., p_{k-1}^*, p_k, p_{k+1}^*, ..., p_N^*) \geq C_l. \end{aligned} \tag{4.18}$$

This fact follows immediately from: 1) the continuity property of the utility function U_k and the interference function Q_l 4.13; 2) the compactness of the simplexes $S_{\varepsilon=0}^{N_k}$; 3) the Nash theorem [79], Th. 1, p.268.

4.2.2 Game as Reinforcement Learning

Now we can formulate the repeated power-control game problem with a *priori* unknown average utility and constraints. The objective of this power-control game is: based on the current information (realized utility and interference), generate the mixed admissible strategy $\{p_k^t\}, \forall k \in N$ to achieve a realizable Nash Equilibrium within the subclass of stationary strategies.

To achieve this objective, we need to emphasize the following fact at first. According to the Nash theorem [79], the set of stationary distribution $(p_1^*, ..., p_N^*)$ that satisfies (4.18) might contain more than one element. For one of the Nash Equilibrium distribution $(p_1^*, ..., p_N^*)$, the objective is to show if an admissible mixed strategy $W_k, \forall k \in N$ converges to such distribution $(p_1^*, ..., p_N^*)$, then the associated random function of both utility and interference Φ_k^t (4.7)and Ψ_l^t (4.9) also converge to the corresponding expected stationary values $U_k(p_1^*, ..., p_N^*)$ and $Q_l(p_1^*, ..., p_N^*)$, respectively. This result is summarized in the following theorem.

Theorem 1: Under the *assumption A1* and *A2*, there exists a strategy $W_k, \forall k \in N$ that is asymptotically stationary realizing a Nash equilibrium:

$$\limsup_{n\to\infty} E\left\{\frac{1}{t}\sum_{\tau=1}^{t}\sum_{k=1}^{N}\|p_k^t - p_k^*\|\right\} \overset{a.s.}{=} 0. \tag{4.19}$$

Then,

$$\limsup_{n\to\infty} E\left\{\frac{1}{t}\sum_{\tau=1}^{t}\sum_{k=1}^{N}|\Phi_k^t - U_k(p_1^*, ..., p_N^*)|\right\} \overset{a.s.}{=} 0,$$

$$\limsup_{n\to\infty} E\left\{\frac{1}{t}\sum_{\tau=1}^{t}|\Psi_l^t - Q_l(p_1^*, ..., p_N^*)|\right\} \overset{a.s.}{=} 0. \tag{4.20}$$

Proof Let us consider the following sequence,

$$\begin{array}{l} S_1^t = \frac{1}{t}\sum\limits_{\tau=1}^{t}\sum\limits_{k=1}^{N}|\Phi_k^t - U_k\left(p_1^*, ..., p_N^*\right)| \\ S_{l,2}^t = \frac{1}{t}\sum\limits_{\tau=1}^{t}|\Psi_l^t - Q_l^t\left(p_1^*, ..., p_N^*\right)| \end{array}. \tag{4.21}$$

According to the assumption A1, it follows

$$\begin{array}{l} U_k^t\left(p_1, ..., p_N\right) = E\left\{\Phi_k^t/H_{t-1}\right\} \\ Q_l^t\left(p_1, ..., p_N\right) = E\left\{\Psi_l^t/H_{t-1}\right\} \end{array}. \tag{4.22}$$

Then, using the above equality (4.22), we derive the recurrent form of (4.21).

$$\begin{array}{l} S_1^t = \left(1 - \frac{1}{t}\right)S_1^{t-1} + \frac{1}{t}|E\left\{U_k^t\left(p_1^*, ..., p_N^*\right) - \Phi_k^t/H_{t-1}\right\}| \\ S_{l,2}^t = \left(1 - \frac{1}{t}\right)S_{l,2}^{t-1} + \frac{1}{t}|E\left\{Q_l^t\left(p_1^*, ..., p_N^*\right) - \Psi_l^t/H_{t-1}\right\}| \end{array}.$$

Let us consider S_1^t first. Take the conditional expectation of $\left(S_1^t\right)^2$, and in the view of the assumption A2, we derive the following equation,

$$E\left\{\left(S_1^t\right)^2|H_{t-1}\right\} = \left(1-\tfrac{1}{t}\right)^2\left(S_1^{t-1}\right)^2$$

$$+\tfrac{2}{t}\left(1-\tfrac{1}{t}\right)E\left\{S_1^t|U_k^t\left(p_1^*,...,p_N^*\right)-\Phi_k^t/H_{t-1}|/H_{t-1}\right\}\ldots(a)$$

$$+\tfrac{1}{t^2}E\left\{|U_k^t\left(p_1^*,...,p_N^*\right)-\Phi_k^t/H_{t-1}|^2/H_{t-1}\right\}\ldots\ldots\ldots\ldots(b)$$

In part (a) of the above equation, according to *Cauchy-Schwarz inequality*, we have

$$E\left\{S_1^t|U_k^t\left(p_1^*,...,p_N^*\right)-\Phi_k^t/H_{t-1}|/H_{t-1}\right\}$$
$$\leq\sqrt{E\left\{\left(S_1^t\right)^2/H_{t-1}\right\}}\sqrt{E\left\{\left(|U_k^t\left(p_1^*,...,p_N^*\right)-\Phi_k^t/H_{t-1}|\right)^2\right\}}.$$

Since the transmission rate function $U_k^t\left(p_1^*,...,p_N^*\right)$ are smooth enough, and it can be checked that U_k^t satisfies the following Lipschitz continuity condition.

$$\sqrt{E\left\{\left(|U_k^t\left(p_1^*,...,p_N^*\right)-\Phi_k^t/H_{t-1}|\right)^2\right\}}\leq\beta\sum_{k=1}^{N}\left\|p_k^t-p_k^*\right\|,$$

where β is a constant. According to the assumptions A1 and A2, $E\left\{\left(S_1^t\right)^2/H_{t-1}\right\}$ must be bounded by a constant value. In part (b), in the view of the assumption A2, we have

$$|U_k^t\left(p_1^*,...,p_N^*\right)-\Phi_k^t/H_{t-1}|^2\leq\sigma.$$

So, accordingly, we have the following inequality for the conditional expectation of $\left(S_1^t\right)^2$,

$$E\left\{\left(S_1^t\right)^2|H_{t-1}\right\}\leq\left(1-\frac{1}{t}\right)\left(S_1^{t-1}\right)^2+\frac{C_1}{t}\left\|p_k^t-p_k^*\right\|+\frac{C_2}{t^2}.$$

By using the similar procedure, we can obtain the same inequality for the interference power part, $S_{l,2}^t$.

In the view of lemma A. 11 [80], this inequality leads to

$$S_1^t\underset{t\to\infty}{\overset{a.s.}{\to}}0,\qquad S_{l,2}^t\underset{t\to\infty}{\overset{a.s.}{\to}}0.\tag{4.23}$$

Using lemma A. 13 [80] and considering the mathematical expectation of both sides of the previous inequality, we obtain the desired results (4.20).

Remark 2: This result shows that in the wireless environment, which is "asymptotically stationary in the average", the convergence of stationary Nash equilibrium distribution of $(p_1^*, ..., p_N^*)$ always guarantee the convergence of transmission rates and interference functions. The proof of this theorem relies on the application of Robbins-Siegmund theorem (Robbins and Siegmund 1971) [81] in probability theory, a convergence theorem for dependent random sequences.

Therefore, if we can construct an asymptotically (with time) stationary mixed strategy $\{p_k^t\}$ converging to a stationary distribution $(p_1^*, ..., \mathrm{p}_\mathrm{N}^*)$ that realizes a Nash equilibrium, we will achieve the main goal of this learning-based repeated power-control game. However, as we known from the game theory, the set of stationary distributions $(p_1^*, ..., \mathrm{p}_\mathrm{N}^*)$ that satisfying (4.17) might not exist, or unique. To attain our main goal in a rigorous manner, at first, the problem related to the existence and uniqueness of the Nash equilibrium should be solved. In the next section, we deal with these problems.

4.3 Properties of Nash Equilibrium

4.3.1 Existence of Nash Equilibrium

To justify the correctness of Lagrange multipliers implementation and existence of nonempty set of admissible transmit power strategies, assume that the given interference constraints satisfy the slater's condition providing the Kuhn-Tucker constraint qualification [82].

Assumption A3: There exists a *feasible* strategy vectors $(\tilde{p}_1^t, ..., \tilde{p}_N^t)$ such that $1 \leq k \leq N$ and $1 \leq l \leq M$, the following inequalities hold:

$$Q_l(\tilde{p}_1^t, ..., \tilde{p}_N^t) \leq C_l. \tag{4.24}$$

The next theorem shows the main result on the existence of Nash Equilibrium of the power-control game.

Theorem 2: An Nash equilibrium point (that satisfies (4.17)) strategy exists for any N-person game satisfying **(A1)** and **(A3)**.

Proof It follows straight from Th.1 in [83]. Indeed, by **(A1)** the function

$$\begin{aligned}\rho(p,q) &= \sum_{k=1}^{N} U_k(p_1,...,q_k,...,p_N) \\ &= \sum_{k=1}^{N} W\log_2(1+\frac{K\mathsf{P}_k\cdot q_k G_{kk}(s_k)}{\sum_{j\neq k}^{j\in N}\mathsf{P}_j\cdot p_j G_{kj}(s_j)+n_k'})\end{aligned} \tag{4.25}$$

is well defined for any $p = (p_1, ..., p_N)$ and $q = (q_1, ..., q_N)$ from $R = (R^1 \times ... \times R^N)$. It is easy to prove $\rho(p, q)$ is continuous in p and q. We can further prove $\rho(p, q)$ is a strictly concave function in q for any fixed p.

Taking the derivative of $\rho(p, q)$ with respect to q_k, we have have

$$\nabla_{q_k}\rho(p,q) = \frac{W\left(K\mathsf{P}_k G_{kk}(s_k)\right)}{\sum_{j\neq \mathrm{k}}^{j\in N}\mathsf{P}_j\cdot \mathrm{p}_j G_{kj}(s_j)+\mathrm{n}_k+K\mathsf{P}_k\cdot \mathrm{q}_k G_{kk}(s_k)} > 0. \tag{4.26}$$

Taking the derivative of the above equation again, we have

$$\nabla^2_{q_k}\rho(p,q) = \frac{-W\left(K\mathsf{P}_k G_{kk}(s_k)\right)^2}{\left(\sum_{j\neq \mathrm{k}}^{j\in N}\mathsf{P}_j\cdot \mathrm{p}_j G_{kj}(s_j)+\mathrm{n}_k+K\mathsf{P}_k\cdot \mathrm{q}_k G_{kk}(s_k)\right)^2} < 0. \tag{4.27}$$

We obtain the Hessian:

$$\mathbf{H}_k = diag(\nabla^2_{q_1}\rho(p,q), ..., \nabla^2_{q_N}\rho(p,q)). \tag{4.28}$$

Matrix $\mathbf{H}_k$ is obviously *negative* definite: for all vectors v,

$$\mathbf{v}^T\mathbf{H}_k\mathbf{v} = diag(\nabla^2_{q_1}\rho(p,q)v_1^2, ..., \nabla^2_{q_N}\rho(p,q)v_N^2). \tag{4.29}$$

Therefore, $\rho(p, q)$ is a strictly concave function in q for any fixed p within R. Notice that R is convex, compact and nonempty by **(A3)** and (4.24). Then, the point-to-set mapping $p \in R \to \Gamma p \subset R$ given by

$$\Gamma_p = \left\{q|\rho(p,q) = \max_{z\in R}\rho(p,z)\right\} \tag{4.30}$$

is upper semi-continuous in R, and, hence, by Kakutani fixed point theorem (see, e.g. [84]), there exists a point $p^* \in R$ such that $p^* \in \Gamma p^*$, that is

$$\rho(p^*,\ p^*) = \max_{z\in R}\rho(p^*,z). \tag{4.31}$$

This fixed point $p^* \in R$ satisfies (4.17). Indeed, suppose that (4.17) is not verified, e.g., for $k = k_0$, there exists a point p_{k_0} such that $p^*(k_0) = (p_1^*, ..., p_{k_0}, ..., p_N^*) \in R$. We have $U_k(p_1^*, ..., p_{k_0}, ..., p_N^*) > U_k(p_1^*, ..., p_k^*, ..., p_N^*)$ that implies $\rho(p^*, p^*(k_0)) > \rho(p^*, p^*)$ which contradicts (4.31).

4.3.2 Uniqueness of Nash Equilibria

The condition for uniqueness of equilibria are known as *strict diagonal concavity*. According to Rosen's theorem 2 in [83], we show that if the given matrix game is "strictly diagonal concave", then the corresponding game turns out to be strictly diagonal concave, and the uniqueness of the equilibria policy follows.

4.3.2.1 Diagonal Concavity Properties

Let us define the function

$$W_r(p, q) = (q - p)\frac{\partial}{\partial p}\rho_r(p, q) + (p - q)\frac{\partial}{\partial q}\rho_r(p, q), \tag{4.32}$$

where $\rho_r(p, q) = \sum_{k=1}^{N} r_k U_k(p_1, ..., q_k...., p_N)$.

Definition 6: A matrix game is said to be *diagonal concave* if there exists positive numbers r_k such that for any $p, q \in R$

$$W_r(p, q) \geq 0; \tag{4.33}$$

a matrix game is said to be *strict diagonal concave* if there is a positive numbers r_k such that for any $p, q \in R$, the strict inequality holds above. The basic properties of the diagonal concave games can be found in [85].

4.3.2.2 Uniqueness of the Power-Control Game

As already mentioned above, the equilibrium probability distribution may not be unique. We will prove that the proposed matrix game for CR power control is strict diagonal concave, and admit a unique Nash equilibrium

Lemma 2: The matrix game for CR user power control in the *interference non-dominance* condition is a strict diagonal concave game, i.e.,

$$W_r(p,q) > 0. \tag{4.34}$$

Proof Let us first expand the function $W_r(p,q)$,

$$\begin{aligned} W_r(p,q) &= (q-p)\frac{\partial}{\partial p}\rho_r(p,q) + (p-q)\frac{\partial}{\partial q}\rho_r(p,q) \\ &= \frac{Wr\cdot(q-p)\left(KP_kG_{kk}(s_k)\right)}{\sum_{j\neq k}^{j\in N} P_j\cdot q_jG_{kj}(s_j)+n_k+KP_k\cdot p_kG_{kk}(s_k)} + \\ &\quad \frac{Wr\cdot(p-q)\left(KP_kG_{kk}(s_k)\right)}{\sum_{j\neq k}^{j\in N} P_j\cdot p_jG_{kj}(s_j)+n_k+KP_k\cdot q_kG_{kk}(s_k)} \\ &= \frac{Wr\cdot(p-q)^2\left(KP_kG_{kk}(s_k)\right)}{\left(\sum_{j\neq k}^{j\in N} P_j\cdot q_jG_{kj}(s_j)+n_k+KP_k\cdot p_kG_{kk}(s_k)\right)} \\ &\quad \cdot\frac{\left(KP_kG_{kk}(s_k)-\sum_{j\neq k}^{j\in N} P_jG_{kj}(s_j)\right)}{\left(\sum_{j\neq k}^{j\in N} P_j\cdot p_jG_{kj}(s_j)+n_k+KP_k\cdot q_kG_{kk}(s_k)\right)}, \end{aligned} \tag{4.35}$$

where $r = (r_1, ..., r_k, ..., r_N)$. Since we have $G_{kk}(s_k) \gg G_{kj}(s_j)$ in the interference non-dominance condition, and K approaches to 1 in a median SINR value of transmissions, it can guarantee that $KP_kG_{kk}(s_k) - \sum_{j\neq k}^{j\in N} P_jG_{kj}(s_j)$ in (4.35) are always positive for the CR users in the most typical weak interference high data rate (SINR) scenarios. So, that proves there must exist a r_k such that $W_r(p,q) > 0$.

Theorem 3: The matrix game for CR user power control admits a unique Nash Equilibrium.

Proof The proof is straight forward, because of the power-control game is strict diagonal concave according to lemma 2. It is a direct outcome of Rosen's theorem [83].

Therefore, we can claim that the power-control game converges to a unique equilibrium point, during the learning process.

4.3.3 Lagrange Multipliers Using Regularization Approach

To get the Nash equilibrium of the CR users' power control strategy in (4.17), Lagrange multiplier method can be used. The corresponding Lagrange function is given by

$$L_k(p_1, ..., p_N; \lambda) = U_k(p_1, ..., p_N) - \sum_{l=1}^{M} \lambda_l \left(Q_l(p_1, ..., p_N) - C_l\right),$$

where $\lambda_l > 0$ and the arguments p_k belong to the $\varepsilon-$simplexes (4.14). Since the interference power constraint (4.15) is not strictly convex (multilinearity of interference function) and, as a consequence, any attempt to directly apply the gradient technique for finding its saddle point may fail because of divergence.

One approach for avoiding this problem consists of introducing a regularization term in the corresponding Lagrange function [86]

$$L_k^{\delta}(\mathrm{p}_1, ..., \mathrm{p}_N; \lambda) = L_k(\mathrm{p}_1, ..., \mathrm{p}_N; \lambda) - \frac{\delta}{2}\left(\|p_k\|^2 - \|\lambda_k\|^2\right), \tag{4.36}$$

where $\lambda_k = (\lambda_{k,1}, ..., \lambda_{k,M})^T$. These regularized functions are strictly concave with respect to p_k, and strictly convex with respect to r_k. The next theorem describes the dependence of equilibrium-point (or saddle-point) $(p_k^*(\varepsilon_k^t, \delta^t), \lambda_k^*(\varepsilon^t, \delta^t))$ of the regularized function with the regularizing $\delta_t, \varepsilon_k^t$ for convergence.

Theorem 4: If the sequence $\{\varepsilon_k^t\}$ and $\{\delta_t\}$ are such that

$$\varepsilon_k^t \in (0, 1/N_k), \delta^t > 0, \lim_{t\to\infty} \delta^t = 0, \varepsilon^t = (\varepsilon_1^t, ..., \varepsilon_N^t)^T$$
$$\lim_{n\to\infty} \varepsilon^t/\delta^t = \upsilon = (\upsilon_1, ..., \upsilon_N)^T, \upsilon_k \in [0, \infty), n = 1, 2, ...$$

the sequence $(p_k^*(\varepsilon_k^t, \delta^t), \lambda_k^*(\varepsilon^t, \delta^t))$ of the equilibrium-points of the corresponding Lagrange function $L_k^{\delta}(\mathrm{p}_1, ..., \mathrm{p}_N; \lambda)$ (4.36) converges to the equilibrium-point $(p_k^{**}, \lambda_k^{**})$ of the initial Lagrange function $L_k(\mathrm{p}_1, ..., \mathrm{p}_N; \lambda)$, which corresponds to the equilibrium-point (p_k^*, λ_k^*) and has the minimal norm, i,e.,

$$(p_k^*(\varepsilon_k^t, \delta^t), \lambda_k^*(\varepsilon_k^t, \delta^t)) \to (p_k^*, \lambda_k^*), t \to \infty. \tag{4.37}$$

where

$$p_k^{**} = \arg\min_{p^*, \lambda^*} \frac{1}{2}(\|p^*\|^2 + \|\lambda^*\|^2). \tag{4.38}$$

Proof The proof is similar to [87], we omit it due to space limitation.

The following two lemmas are two useful conclusions related to Theorem 4, which will be used in the convergence analysis of the learning algorithm in Section VI.

Lemma 3 For any $p \in P_N$ and for any $\lambda \in R^M$ the following inequality holds:

$$\begin{aligned}
&-\tfrac{\delta}{2}\sum_{k=1}^{N} r_k\left(\left\|p_k^t - p^*(\varepsilon_k^t,\delta^t)\right\|^2 + \left\|\lambda_k^t - \lambda^*(\varepsilon_k^t,\delta^t)\right\|^2\right)\\
&\le \sum_{k=1}^{N} r_k[(p_k^t - p^*(\varepsilon_k^t,\delta^t))^T \tfrac{\partial}{\partial p_k^t} L_k^\delta(p_1,...,p_N;\lambda_n)\\
&-(\lambda_k^t - \lambda^*(\varepsilon_k^t,\delta^t))^T \tfrac{\partial}{\partial \lambda_n} L_k^\delta(p_1,...,p_N;\lambda_n)] \le 0.
\end{aligned}$$

where (p_k^*, λ_k^*) is the equilibrium-point of the regularized Lagrange function.

Proof The right-hand-side inequality is always hold due to the Kuhn-Tucker saddle point theorem [86]; the right-hand-side inequality is obtained according to the property of Kuhn-Tucker saddle point theorem and the Jensen's inequality [86]. The detailed proof is omitted due to space limitation.

Lemma 4: Under the same conditions as in Theorem 4, 1) all the possible equilibrium mixed policies can be parameterized by the nonnegative vector parameter $\upsilon \in R^N$ as $\left(p_1^*(\upsilon),...,p_N^*(\upsilon)\right)$ and the sequences $(p_1^*(\varepsilon_1^t,\delta^t),...,p_N^*(\varepsilon_N^t,\delta^t))$ $\lambda_k^*(\varepsilon^t,\delta^t)$ to the unique equilibrium points $\left(p_1^*(\upsilon),...,p_N^*(\upsilon)\right)$ and $\lambda_k^*(\upsilon)$, i.e.,

$$\sum_{k=1}^{N} r_k\left(\left\|p_k(\upsilon) - p^*(\varepsilon_k^t,\delta^t)\right\|^2 + \left\|\lambda_k(\upsilon) - \lambda^*(\varepsilon_k^t,\delta^t)\right\|^2\right) \underset{n\to\infty}{\to} 0.$$

2) There exists constants $c_i (i = 1,2,3)$ such that

$$\begin{aligned}
&\sum_{k=1}^{N} r_k\left(\left\|p_k^*(\varepsilon_k^{t+1},\delta^{t+1}) - p_k^*(\varepsilon_k^t,\delta^t)\right\|^2 +\right.\\
&\left.\left\|\lambda_k^*(\varepsilon^{t+1},\delta^{t+1}) - \lambda_k^*(\varepsilon^t,\delta^t)\right\|^2\right)\\
&\le c_1\left\|\varepsilon^{t+1} - \varepsilon^t\right\| +\\
&c_2\left\|\delta^{t+1} - \delta^t\right\| + c_3\left\|\varepsilon^{t+1}/\delta^{t+1} - \varepsilon^t/\delta^t\right\| = \beta_t.
\end{aligned} \tag{4.39}$$

Proof A similar proof can be found in [87]. The detailed proof is omitted due to space limitation.

4.4 Reinforcement Learning Algorithm

4.4.1 The Complete Information Case

At first, let us review the power control problem in the complete information case, when the expected utility and interference functions[4] are available. In this case, the Lagrange multiplier and the projection of gradient procedure can be applied for the class of diagonal concave games to attain the equilibrium point (see a work [24])

$$\begin{cases} p_k^{t+1} = p_k^t + \gamma_k^t \nabla_{p_k} L_k^{\delta_t}\left(p_1^t, ..., p_N^t; \lambda^t\right) \\ \lambda_{k,l}^{t+1} = \lambda_l^t - \gamma_k^t \nabla_{\lambda_l} L_k^{\delta_t}\left(p_1^t, ..., p_N^t; \lambda^t\right) \\ p_k^{t+1} = \pi_{S_{\varepsilon=0}^{N_k}}\left(p_k^{t+1}\right), \lambda_l^{t+1} = \left[\lambda_l^{t+1}\right]^+ \\ \forall k = 1, .., N, t = 1, 2, ... \end{cases}, \tag{4.40}$$

where $\pi_{S_{\varepsilon^t=0}^{N_k}}\{\cdot\}$ is the projection operation to the simplex $S_{\varepsilon^t=0}^{N_k}$, λ_l^t is the dual price for η_l^t, and $[\cdot]^+$ is the "take positive part" operator. If the parameters of this procedure satisfy $\gamma_k^t, \delta_t, \varepsilon^t \to 0$ and $\sum_{t=1}^{\infty} \gamma_k^t \delta^t = \infty, \sum_{t=1}^{\infty} \beta^t < \infty$. ($\beta^t$ is defined by (4.39)), it provides the convergence of p_k^t to the unique equilibrium strategy $\left(p_1^*(\upsilon), ..., p_N^*(\upsilon)\right)$ in the Lemma 4.

4.4.2 The Incomplete Information Case

In the case of incomplete information case that concerned in this work, that is, only the current realizations of utility u_k^t and the function of interference level η_l^t explored from PUs are available; the direct application of the Lagrange multiplier approach is infeasible. The "stochastic approximation" version of (4.40) can be applied instead of $\nabla_{p_k} L_k^{\delta_t}\left(p_1^t, ..., p_N^t; \lambda^t\right)$ and $\nabla_{\lambda_l} L_k^{\delta_t}\left(p_1^t, ..., p_N^t; \lambda^t\right)$. This new procedure requires the estimation of the mixed power control strategies by using the implementation of the current realizations. This procedure is known as *reinforcement leaning algorithm*, which can be implemented in different ways. In order to present concrete learning procedures, Let us first briefly introduce the basic concept of reinforcement schemes in learning automation [80].

[4] more specifically, the interference function of every CR user k imposed on PU l, $\eta_{k,l}^t$ are known for each CR user in the networks. Thus, the sum-interference η_l^t is known in advance.

4.4.3 Reinforcement Schemes in Learning Automation

It has been shown in Section II lemma 1 that the stochastic constrained power-control game in (4.7),(4.9) is asymptotically equivalent to the problem related to the determination of the equilibrium-point of the regularized function $L_k^{\delta}(p_1, ..., p_N; \lambda)$, using the realizations of the cost function and the constraints. This equivalent problem can be formulated and solved as the behavior of a variable-stochastic automation in multi-teacher environment [88]. Fig. 2 refers to the schematic block diagram for the learning automation to operating in a multi-teacher environment. We note that the normalized procedure processes as a mapping from the teachers' responses $(\zeta_k^t, 1 \leq k \leq N)$ to the learning automation input of CR users' utility (u^t) and (η^t),$1 \leq k \leq N, 1 \leq l \leq M$. The role of the environment is to establish the relation between the actions of the automation and the signals received at its input, which is described as the learning automation in Section I.

Reinforcement schemes are found successful application in the field of learning automata [62]. A reinforcement scheme is similar to the recursive estimation procedure used in adaptive control. The reinforcement scheme generates p^{t+1} from p^t based on incremental changes in the probabilities. Several algorithms for adjusting the probabilities after each sampling period (interaction with the environment) have been proposed [87]. The most commonly used one is a linear updating algorithm have been proposed by Bush and Mosteller [62]. All the reinforcement schemes described in the literature can be considered as being solutions of optimization problems. The following describe the average penalty function for a single user i in the repeated power-control game.

$$J = \{\Phi_i(p^t)E_i[1 - \xi_i] - \Psi_i(p^t)E_i[\xi_i]\}. \tag{4.41}$$

where the functions $\Phi_i(p^t)$ and $\Psi_i(p^t)$ represent the amount of change in the probability vector under the expected reward ($\xi_i = 0$) and expected penalty ($\xi_i = 1$). In the case when the complete information on the expected utilities and constraints (i.e., the channel and mixed power strategy information for CR users and/or PUs) is available, then the gradient-like technique can be applied for the class of diagonal concave power-control games to

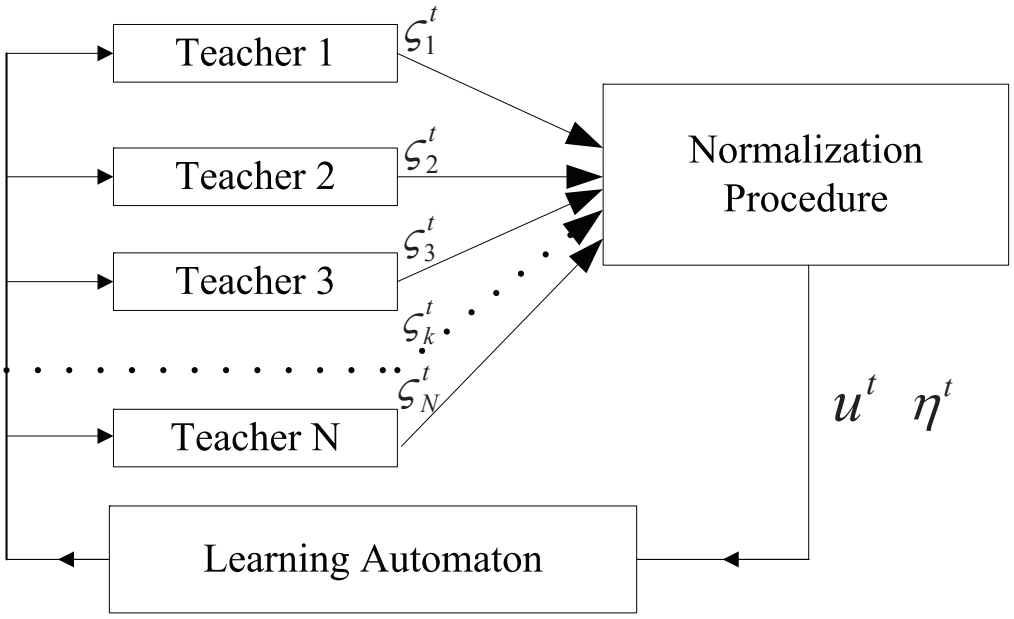

Figure 4.2. Learning automaton of multi-teacher environment

attain the equilibrium point. This corresponds to and similar to the traditional power control schemes in CR networks as [26] [24].

To minimize the penalty function J, the reinforcement scheme that sets the gradient of J equal to zero is derived. At the time step t, the following algorithm is obtained:

$$p_i^{t+1} = p_i^t + \frac{\gamma^t}{p_i^t}\left[\frac{\partial \Phi_i(p^t)}{\partial p_i^t}(1-\xi_i) - \frac{\partial \Psi_i(p^t)}{\partial p_i^t}\xi_i\right]. \tag{4.42}$$

We have just derived is one of the central result of reinforcement schemes. In general, all the existing learning schemes fall into the general framework (4.42) in the learning-based game-theoretical problem setting up.

In this work, the widely used Bush-Mosteller reinforcement scheme [62] is adopted, which is a linear reinforcement scheme. And, the function $\Phi(p^t)$ is $p_i^t - \frac{1}{2}\left(p_i^t\right)^2$ and function $\Psi(p^t)$ is $\left(p_i^t\right)^2$, respectively.

4.4.4 Bush-Mosteller-Reinforcement-based Lagrange Multipliers

For the unknown expected utility and constraints of each CR user, the Bush-Mosteller reinforcement scheme and the normalization procedure presented in [87] will be applied

hereafter to design a new learning algorithm for the N-user constrained repeated power-control game. In fact, we assume that after each stage, the utility as well as the constraints to each CR user are random variables. No information concerning the distribution of the utility and constraints is available. The necessary information is obtained during the course of the game. The learning control is an iterative process involving an adaption at each stage (or time).

Since our formulated power control problem in (4.7) and (4.9) is a *stochastic optimization problem* on discrete sets using only the corresponding observations of utility $\mathrm{u}_k^t(p_k^t)$ and interference $\eta_l^t(p_k^t)$ to be optimized. It is evident that we can not directly use these schemes for the optimization purposes. In fact, environment responses (in the probability sense) have to belong to the unit interval $[0, 1]$. But the available observations do not obligatory satisfies this conditions.

In order to solve this problem, a procedure called "normalization procedure" (as described in Fig.2), which establish the connection between the environment response ζ_k^t and the available observations (u^t) and (η^t) has been initially described in [80] in order to keep the automation input within the unit segment. For the normalized procedure of the Bush-Mosteller scheme [62], we have the following lemma from [87], Lemma 2, p.268:

Lemma 5: If, some positive monotonically decreasing $\{\tau_k^t\}$ $(0 < \tau_k^t \downarrow 0)$ $\forall k \in N$, and positive increasing $\{\lambda^+\}$ $(0 < \lambda^+ \uparrow \infty)$ sequences, the parameters a_k^t and b_k^t are given by

$$\begin{aligned} \mathrm{a}_k^t &= \tilde{a}_k^t(\sigma_{u,k}^+ + \lambda_t^+ \textstyle\sum_{l=1}^M \sigma_{\eta,k,l}^+ + \delta^t)^{-1} \\ \mathrm{b}_k^t &= \tilde{a}_k^t + (N_k - 1)(\tau_k^t)^2(1 + (N_k - 2)\tau_k^t)^{-1} \end{aligned} \tag{4.43}$$

where

$$\begin{aligned} \tilde{a}_k^t &= \tau_k^t(1 - \tau_k^t)(2(1 + (N_k - 2)\tau_k^t))^{-1} \\ \tau_k^{t-1} &= \varepsilon_k^t. \end{aligned} \tag{4.44}$$

then, the automaton input belongs to $(0, 1]$.

Now, the "four-step" recursive algorithm for the learning-based power-control game is presented.

Algorithm: ***Reinforcement-Learning-based CR Power Control***

Step 1: Based on The available data P_k^t=$\mathsf{P}_k(i_k)$, u_k^t, $\eta_{k,l}^t$, $p_k^t(p_k^t(i_k) \geq \varepsilon_k^t > 0)$, λ_k^t build the following functions:

$$\tilde{u}_k^t = u_k^t - \delta^t p_k^t(i_k) - \sum\nolimits_{l=1}^{M} \lambda_l \eta_l^t \tag{4.45}$$

and normalize (scale) them according to the following procedure:

$$\zeta_k^t = 1 - (\mathrm{a}_k^t \tilde{u}_k^t + \mathrm{b}_k^t)/p_k^t(i_k) \tag{4.46}$$

where the sequence $\{a_k^t\}$ and $\{a_k^t\}$ are defined as follows:

$$\begin{aligned} \mathrm{a}_k^t &= \tilde{a}_k^t(\sigma_{u,k}^+ + \lambda_t^+ \textstyle\sum_{l=1}^{M} \sigma_{\eta,k,l}^+ + \delta^t)^{-1} \\ \mathrm{b}_k^t &= \tilde{a}_k^t + (N_k - 1)(\tau_k^t)^2(1 + (N_k - 2)\tau_k^t)^{-1} \\ \tilde{a}_k^t &= \tau_k^t(1 - \tau_k^t)(2(1 + (N_k - 2)\tau_k^t))^{-1} \\ \tau_k^{t-1} &= \varepsilon_k^t. \end{aligned} \tag{4.47}$$

Step 2: Set the initial values of probability distribution among all the transmit power set: $p_k^1(i)$, for $0 \leq i \leq N_k, \sum_{i=1}^{N_k} p_k^1(i) = 1$. Update the probability distribution p_k^{t+1} and the Lagrange multipliers λ^{t+1} using the following iterative schemes:

$$p_k^{t+1} = p_k^t + \gamma_k^t [e_{N_k}(\mathsf{P}_k^t) - p_k^t + \frac{\zeta_k^t(e^{N_k} - N_k e_{N_k}(\mathsf{P}_k^t))}{N_k - 1}] \tag{4.48}$$

where

$$e_{N_k}(\mathsf{P}_k^t) = (\underbrace{0, ..., 0, 1}_{i_k}, 0, ..., 0)^T \in R^{N_k} \tag{4.49}$$

If $\mathsf{P}_k^t = \mathsf{P}_k(i_k)$ and $e^{N_k} = (1, ..., 1)^T \in R^{N_k}$. Here, the time-varying correction (adaptation) factors γ_k^t belong to the unit segment. Notice that the conditional mathematical expectation of the normalized environment responses $\tilde{u}_k^t e_{N_k}(\mathsf{P}_k^t)$ is equal to the gradient of the augmented Lagrange function with respect to the probability distributions, i.e.,

$$\mathbf{E}\{\tilde{u}_k^t e_{N_k}(\mathsf{P}_k^t) | H_{t-1}\} := \frac{\partial}{\partial p_k^t} L_k^\delta(p_1, ..., p_N; \lambda_n). \tag{4.50}$$

The Lagrange multipliers are adjusted according to the following lowing recursion; $\lambda_{k,l}^1 > 0, \gamma_\lambda^t \geq 0$

$$\begin{aligned} \lambda_l^{t+1} &= \left[\lambda_l^t - \gamma_\lambda^t \psi_l^t\right]_0^{\lambda_+^{t+1}} \\ \psi_l^t &= \delta^t \lambda_l^t - \eta_l^t + C_l. \end{aligned} \tag{4.51}$$

Also, note that the conditional mathematical expectation of ψ_l^t is equal to the gradient of the augmented (regularized) Lagrange function with respect to the Lagrange multipliers, that is

$$\mathbf{E}\{\psi_l^t | H_{t-1}\} := \frac{\partial}{\partial \lambda_k^t} L_k^\delta(p_1, ..., p_N; \lambda_n). \tag{4.52}$$

The operator

$$[x]_0^{\lambda_{t+1}^+} = \begin{cases} x, & if x \in [0, \lambda_{t+1}^+] \\ \lambda_{t+1}^+, & if x > \lambda_{t+1}^+ \\ 0, & if x < 0. \end{cases}$$

Step 3: According to

$$\Pr\{\mathsf{P}_k^{t+1} = \mathsf{P}_k(i) | H_t\} = p_k^{t+1}(i) \tag{4.53}$$

generate randomly new discrete random variables p_k^{t+1} for each CR user as in the learning stochastic automata implementation, and get a new observations (realizations) u_k^{t+1} and η_l^{t+1} that correspond to the environment vector-reactions.

Step 4: Return to Step 1.

The positive sequence $\{\varepsilon_k^t\}, \{\delta^t\}, \{\lambda_t^+\}, \{\gamma_k^t\}$, and $\{\gamma_\lambda^t\}$ will be defined next.

This adaptive Learning algorithm is constructed using the Bush-Mosteller reinforcement scheme (4.48) with the time-varying correction factors γ_k^t, continuous input $\tilde{u}_k^t$, and a normalization procedure which is used to ensure the probability measure. It is easy to verify that $\zeta_k^t \in (0, 1)$ for any time t, and $p_k^t \in S_{\varepsilon_n}^{N_k}$.

4.5 Convergence Analysis and Learning Rate

4.5.1 Convergence Analysis

The following theorem acclaims the convergence of the learning-based repeated power-control game.

Theorem 5 Suppose that the *Assumption* A1-A3 hold for the learning reinforcement procedure (4.48)- (4.51) and the CR power control is diagonal concave. In addition, assume that:

- There exists four nonnegative sequences $\{\varepsilon_k^t\}, \{\delta^t\}, \{\gamma_k^t\}$, and $\{\gamma_\lambda^t\}$ such that $\{\gamma_k^t\} \downarrow 0$, $\delta^t \in (0, \delta^+)$, $\delta^t \downarrow 0$ and $\varepsilon_k^t \in (0, (1/N_k))$, $\varepsilon_k^t \downarrow 0$, $limsup_n(\varepsilon_k^t/\delta^t) < \infty$;

- The updating factor γ_k^t, and γ_λ^t are selected as

$$\gamma_{\lambda,k}^t = \frac{\gamma_k^t a_k^t N_k}{N_k - 1}, 1 \le k \le N. \tag{4.54}$$

  where a_k^t is defined by (4.47) and γ_λ^t satisfies $\sum_{t=1}^{\infty} \delta^t \sum_{k=1}^{N} \gamma_{\lambda,k}^t = \infty$;

- The following series converges:

$$\sum_{t=1}^{\infty} \left[\varphi^t + \beta_t^2 \left(\delta^t \sum_{k=1}^{N} \gamma_{\lambda,k}^t \right)^{-1} \right] < \infty. \tag{4.55}$$

  where $\varphi^t = \beta_t^2 + C_\varphi^2 \sum_{k=1}^{N} (\gamma_{\lambda,k}^t)^2 + 2C_\varphi \beta_t \sum_{k=1}^{N} \gamma_{\lambda,k}^t$
  $+ 2M \sum_{k=1}^{N} (\gamma_{\lambda,k}^t)^2 \left[(\delta^t \lambda_t^+)^2 + \left(\Phi_\eta^+\right)^2 \right]$
  $+ 2\sqrt{M} \sum_{k=1}^{N} (\gamma_{\lambda,k}^t)(\delta^t \lambda_t^+ + \Phi_\eta^+)$.

Then, the mixed strategy of CR users ensure the convergence of the game to the equilibrium point, i.e.,

$$\begin{aligned} &\sum_{k=1}^{N} r_k(\left\| p_k^*(\varepsilon_k^{t+1}, \delta^{t+1}) - p_k^*(\varepsilon_k^t, \delta^t) \right\|^2 \\ &+ \left\| \lambda_k^*(\varepsilon^{t+1}, \delta^{t+1}) - \lambda_k^*(\varepsilon^t, \delta^t) \right\|^2) \underset{t\to\infty}{\to} 0. \end{aligned} \tag{4.56}$$

Proof The techniques used in the proof are based on the Lyapunov approach and Martingale's theory. A proof for similar problem can be referred in [87], page 123. We omit the proof here for brevity and not aside from the main subject of our discussion.

Corollary 1: For the class of the algorithm design parameters defined as follows:

$$\begin{aligned} &\gamma_k^t = \gamma_k^0 t^{-\gamma}, \varepsilon_k^t = \varepsilon_k^0 t^{-\varepsilon}, \gamma_k^0, \varepsilon_k^0, \delta^0, \lambda_0^+ > 0 \\ &\delta^t = \delta^0 t^{-\delta}, \lambda_t^+ = \lambda_0^+ + t^\gamma \ln t. \end{aligned} \tag{4.57}$$

The conditions of Theorem 5 will be verified is

$$\gamma + \varepsilon + \delta + \lambda \leq 1 (\gamma > 0, \varepsilon \geq \delta > 0, \lambda \geq 0) \tag{4.58}$$

and the convergence is ensured if

$$2\gamma > 1. \tag{4.59}$$

Proof The proof of the convergence follows directly by substituting (4.57)-(4.59) into the conditions of the previous theorem and in view of the fact that

$$\sum\nolimits_{n=1}^{\infty} t^{-\alpha} \begin{cases} = \infty, if \alpha \leq 1 \\ < \infty, if \alpha > 1. \end{cases}$$

4.5.2 Learning Rate

As we know, not only the convergence of the power-control game is important but the speed is also essential. For the specific class of the design parameters (4.56), the next theorem states the convergence rate of the learning game algorithm described above.

Theorem 6: Under the condition of the previous theorems and for the class of design parameters (4.56), there exists v such that,

$$W_*^t = \sum_{k=1}^{N} r_k \left(\left\| p_k^t - p^*(\upsilon) \right\|^2 + \left\| \lambda_k^t - \lambda^*(\upsilon) \right\|^2 \right) = o(n^{-v}),$$

where the order v of the adaptation rate satisfies the following constraint:

$$v < v^*(\gamma, \delta, \varepsilon, \lambda) \leq v^{**} = \tfrac{1}{3}$$

$$v^*(\gamma, \delta, \varepsilon, \lambda) = \min\{2\gamma - 1; \gamma + \delta; \varepsilon - \delta + \gamma; 2\delta\}$$

and the maximum adaptation rate $v^{**}=v^*(\gamma^*, \delta^*, \varepsilon^*, \lambda^*)$ is reached for

$$\varepsilon=\varepsilon^*=\delta=\delta^*=\frac{1}{6}, \gamma=\gamma^*=\frac{2}{3}, \lambda=\lambda^*=0.$$

Proof The expression of $v^*(\gamma, \delta, \varepsilon, \lambda)$ follows from [87] Lemma A.3-2, App. A. The optimal design parameters are the solution of the following constrained optimization problem

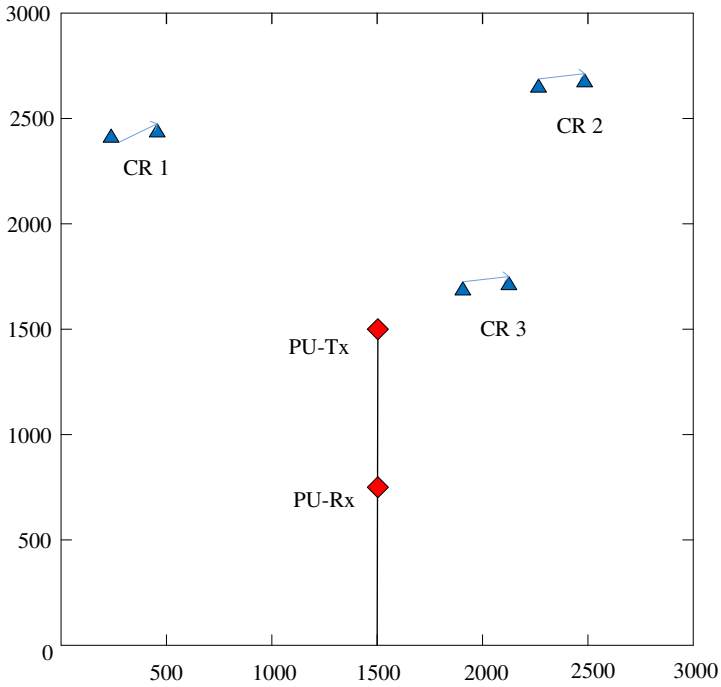

Figure 4.3. User location in the network.

$v^*(\gamma, \delta, \varepsilon, \lambda) \rightarrow max$ over tall the parameters $\varepsilon \geq \delta$ satisfying (4.57) and (4.58). The solution of this problem is achieved when all minimal terms within the operator turn out to be equal, that is

$$\begin{aligned} v^{**} &= \max_{(\gamma,\delta,\varepsilon,\lambda)} \min\{2\gamma - 1; \gamma + \delta; \varepsilon - \delta + \gamma; 2\delta\} \\ &\leq (by\ \gamma \leq 1 - \delta - \varepsilon - \lambda) \\ &\max_{1-\delta-\varepsilon-\lambda} \min\{1 - 2\varepsilon - 2\lambda; 1 - \varepsilon - \lambda; 1 - 2\delta - \lambda; 2\delta\} \\ &= \max_{1-\delta-\varepsilon-\lambda} \min\{1 - 2\varepsilon - 2\lambda; 2\delta\} \\ &= (\text{by}\varepsilon = 1/2 - 2\delta - \lambda) \\ &= \max_{\delta \leq \varepsilon} 2\delta = 1/3, \text{under } \delta = 1/6 \text{ and } \lambda = 0. \end{aligned}$$

4.6 Simulation Results

In this section, simulation results are presented for the repeated power-control game in CR networks. We consider there are multiple CR Tx-Rx pairs and one PU Tx-Rx pair sharing a spectrum band with the bandwidth of $1MHZ$. We set up a system of CR users and the

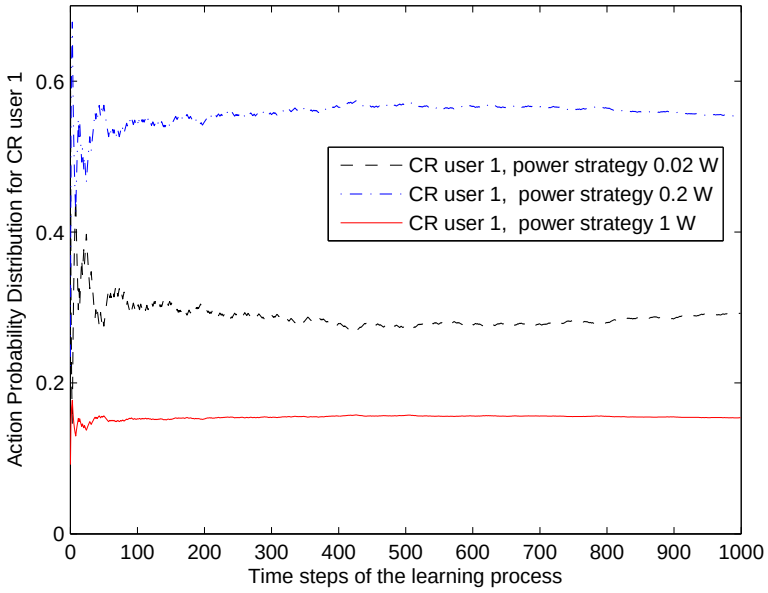

Figure 4.4. Learning process of the mixed strategy: CR user 1.

PU with their locations shown in Fig.3. The simulation parameters are set as: the distance of the CR Tx-Rx is $250m$; The action sets of transmit power is $\mathsf{P}_k = \{0.02, 0.2, 1\}Watt$, the noise of the measurement is according to a zero-mean Gaussian noise, where its power is $\sigma_k^2 = -118.45dBm$. The channel gain is settled to be a log-normal shadowing path loss model, $g_k = (1/d_k)^4$. For the fading part, we use the finite-state Markov channel (FSMC) to model the Rayleigh fading channel, where the number of sates and parameters of the channel model is given in [89]. This network set-up guarantees the *interference non-dominance* condition, where the interference from other CR users and the background noise is less than the transmit power of each CR user. We consider the PU to be always active, where $b_l = 1$. The interference power constraint can change and be set to different values. The default interference power constraint in the PU-Rx is set to be $-115.37dBm$, which is a light interference constraint. As a typical illustration, a random initial probability distribution of the mixed strategies are generated for the CR user 1, the CR user 2 and the CR user 3: $p_1^1 = \{0.59, 0.28, 0.13\}$, $p_2^1 = \{0.74, 0.21, 0.05\}$, and $p_2^1 = \{0.84, 0.07, 0.09\}$, respectively. For the utility function in (3), we set $K = 1$, where the transmission rate is the

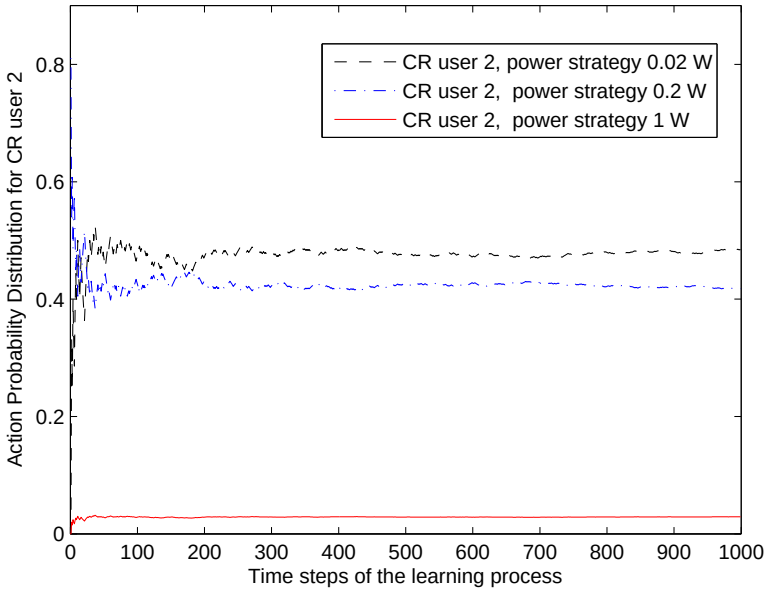

Figure 4.5. Learning process of the mixed strategy: CR user 2.

information-theoretical capacity. The time duration of the learning process for the game is considered in the CRN with the period of $t = 1000$ steps. The parameters of the learning procedure were as follows:

$$\varepsilon = \varepsilon^*=\delta=\delta^*=16, \gamma=\gamma^*=2/3, \varepsilon^0 = 0.25$$
$$\lambda = \lambda^* = 0, \lambda_0^+ = \delta^0 = 0.1, \gamma^0 = 0.15,$$

The proposed learning-based power-control algorithm is implemented based on all initial values give above. And, all the sequences will be decreased according to the learning algorithm until convergence.

Fig. 4, Fig. 5 and Fig. 6 show the convergence process of the probability distribution of the mixed strategy of each CR users. The first observation from our simulation results is that, when every time we generate a random initial probability distributions of the mixed strategies, the equilibrium state of the transmit power level calculated by the mixed strategy of each user is independent with these initial values. This result confirms the theory of convergence in Section VII. Secondly, when comparing the equilibrium state of the mixed strategy of CR users 1, 2, and 3, we can observe that the nearer the CR Tx-Rx to the other

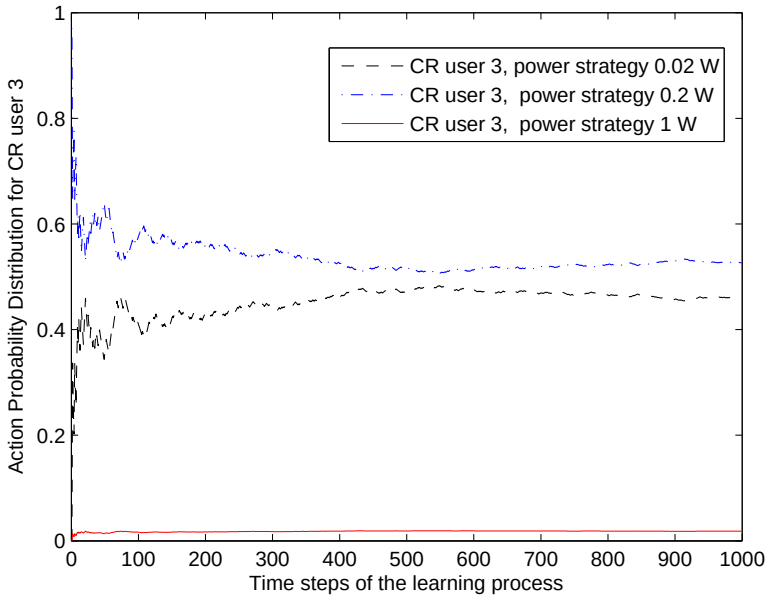

Figure 4.6. Learning process of the mixed strategy: CR user 3.

CR Tx-Rxs, the less the transmit power level (a dot product of mixed strategy probability and the transmit power set) in the equilibrium state. However, we donot observe much more interference penalty on CR users' transmit power, when CR users are nearer to the PU-Rx. An illustration on this point is the CR user 3. Even though the CR user 3 is much more hear to the PU-Rx than the CR user 2 and CR user 1, thus the interference is larger, the transmit power of the CR user 3 is just a little bit higher than the CR user 2. The inside rationale is as follows: as we know, each CR user has blind CSI information from its Tx to the PU-Rx, and CR users can only explore the sum-interference measured by PU-Rx from its feedback control link, the interference penalty price λ_k^t for each CR user in equ. (4.51) is the same for each CR user. Correspondingly, the expected-version of subgradient updating process of the interference penalty in (4.52) that is plugged into equ.(4.40) is the same for each user. However, on the other hand, in equ. (4.45), we can observe the utility function of each CR user is different due to the inter-CR-user interference level, the CR user 1 suffers the least interference, the CR user 3 suffers less, and the CR user 2 suffers the most heavy interference. This affect the expected value of subgradient of transmit power

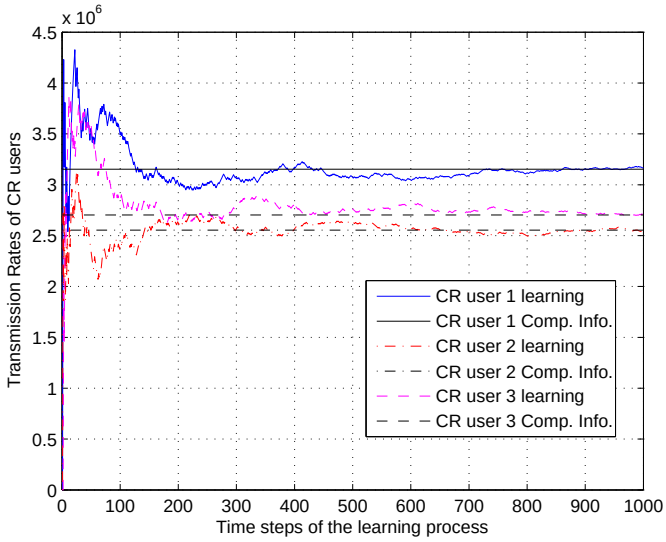

Figure 4.7. Learning process of the transmission rates.

in equ. (4.50), and finally affect the transmit power updating process in equ.(4.40). The result can be clearly viewed from the equ.(4.40): the higher the expected subgradient the higher the equilibrium transmit power. This result is similar to the result in the classic game-theoretical approach without the "pricing" scheme[5]. The simulation results shown in Fig. 4, Fig. 5 and Fig. 6 conform the theoretical analysis.

Fig. 7 shows the learning process of averaged transmission rate for the CR users under the light interference constraints condition of the PU. The results are compared with the theoretical results of the power-control game with complete information case: each CR user knows all the CSI and transmit power strategy in the network, and then the sum-interference of all CR users η_l^t can be calculated by each CR user in the power control process in equ.(4.40). This scenario is equivalent to the classic power-control game case in CR networks without the pricing schemes from PUs as shown in [26], [24], [27]. It can be observed that the averaged transmission rate of each CR user in the learning process will converge and approach to the equilibrium point in the complete information game case, and this simulation results confirm the conclusion of the Lemma 1. As we can see, at this

[5]There is no incentive of the interference power penalty for each individual CR user from PUs based on the interference level of each individual CR user.

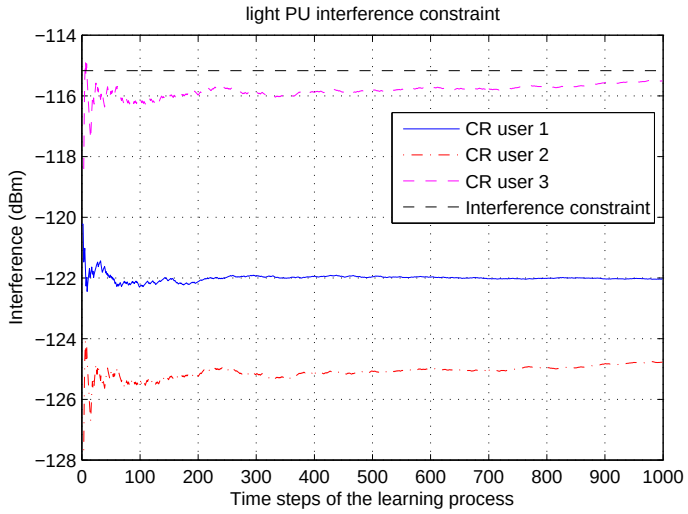

Figure 4.8. Learning process of the interference level: light.

time, the time interval (or iterations) required to achieve the equilibriums is longer than the classic one. This is the expense of no prior knowledge of CSI and transmit power strategy of other users in the network. However, in this scenario, CR users only need to explore several bits information of the interference-level from the feedback link of the PU and use the transmission rate they tried in each step, where the algorithm implementation complexity is very low. It is suitable and adaptable for highly dynamic changing radio-access environments.

Fig.8 shows the convergence process of the interference power of each CR user in the light interference constraint condition. In this case, the CR users transmit with higher power in the equilibrium condition than the CR users in the heavy interference constraint condition (-117.47dBm) in Fig. 9. An interesting observation is, when the interference constraint is low (as shown in Fig. 9), the convergence process of interference power of CR users change much more rapidly in the initial steps than the case when the interference constraint is high (as shown in Fig.8). Especially, this phenomenon is clearly seen for the dominant interference CR users, e.g., the CR user 3 in Fig. 8 and Fig. 9. Moreover, it can be found that, at the initial stages of the repeated power-control game, the interference power constraint of the PU is violated, and it is gradually being met during the learning

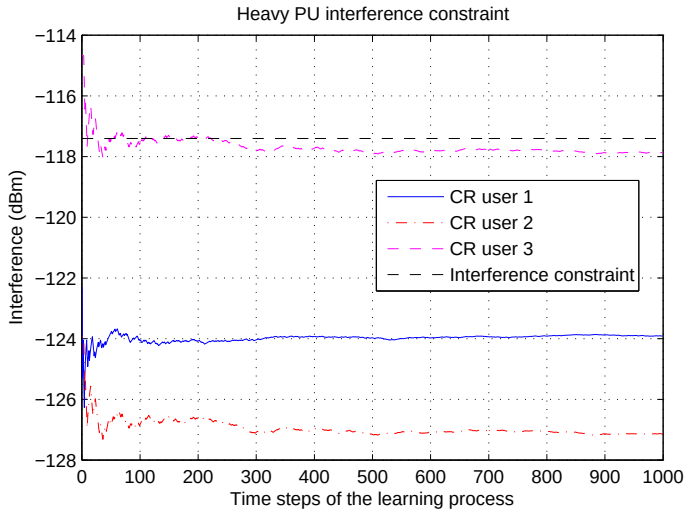

Figure 4.9. Learning process of the interference level: heavy.

process. This is consistent with the theoretical predictions in (29) (32) (33): the learning algorithm will enforce the transmit power strategy largely to reduce the interference caused by each CR user until it conforms to the interference power constraint.

CHAPTER 5

CAPACITY AND DELAY SCALING IN COGNITIVE RADIO AD HOC NETWORKS: IMPACT OF PRIMARY USER ACTIVITY

5.1 Network Model and Definitions

In this section, we describe the network model and define the throughput capacity and delay in CRAHNs.

A. Network Architecture

The network consists of two components: CRAHN and primary network. CRAHN is an ad hoc network containing only secondary nodes. As in [8] [12][13][14], we assume a random network model with static secondary nodes for CRAHNs. Consider a square of unit area $A = [0, 1]\times[0, 1]$, in which secondary nodes are randomly distributed according to a Poisson point process (P.P.P.) of density n. Let the transmission range of each secondary node denote by $r(n)$. As indicated in [90], the transmission range $r(n)$ needs to satisfy $r(n) \geq \Theta(\sqrt{\log n/n})$ to make the network connected. The source-destination (S-D) pairs are randomly chosen in the network region.

In addition to the CRAHN, there are m pairs of independent primary nodes distributed in the unit area square. Each pair of primary nodes contains a primary transmitter and its intended receiver. For spectrum overlay scenario, the transmission range $r(m)$ of these primary nodes should have no intersection, otherwise interference will occur. To ease the analysis of the problem, we assume each primary receiver's *interference range* is within the primary nodes' transmission range to avoid interfering secondary transmitters that lie outside but near the primary nodes' transmission range [1]. This assumption accords to the typical channel access model for two heterogenous networks: the symmetric channel access model as in [91] [92], where a channel can be used when neither primary

[1]For general cases as in [91], the interference range doesn't need to be totally covered in the transmission range, but this shape difference doesn't affect the scalability of CRAHNs while our analysis is largely simpled.

nodes' interference range are within secondary nodes' transmission range nor secondary nodes' interference range are within primary nodes' transmission range. It guarantees the interference-free transmission of both forward and backward links of primary networks and CRAHN. This scenario corresponds to (but not limited to) a broadcast network, such as the TV or the cellular networks, in which the primary transmitters are fixed base-stations.

For spectrum overlay, when any primary transmitter is active, all the traffic of secondary nodes in its transmission range cannot be transmitted, or transmission collision occurs. To derive the achievable throughput capacity for CRAHN, we assume secondary nodes can always utilize the off time of primary transmissions, and there are no sensing errors and the switching delay between the on and off states. In this fundamental work, PU activity can be modeled as exponentially distributed inter-arrivals [93] two-state birth-death process, and it is i.i.d for each pair of primary nodes. The birth rate (active periods) has a geometric distribution with parameter p_B and the death rate (inactive periods) has a geometric distribution with parameter p_I. The average busy and idle times are thus given by $\alpha = 1/p_B$ and $\beta = 1/p_I$. An ON (active) state represents the period used by PUs, and an OFF (inactive) state represents the unused period. Since each user arrival is independent, each transmission follows the Poisson arrival process. So, we can estimate the probabilities of ON and OFF states as follows:

$$P_{on} = \frac{\alpha}{\alpha + \beta}; \quad P_{off} = \frac{\beta}{\alpha + \beta}.$$

B. Interference Model of CRAHN

The interference model used in this work are the protocol model for dense networks introduced in [8], which is suitable for overlay CRAHNs: a transmission from node X_i is successfully received by node X_j if the following two conditions are satisfied: 1)$|X_i - X_j| \leq r(n)$ and 2)$|X_k - X_j| \geq (1+\Delta)r(n)$. X_i, X_j and X_k are the locations of secondary nodes. $\Delta > 0$ guarantees a guard zone to prevent a neighboring secondary node from transmitting on the same channel simultaneously.

C. Routing Strategies of the CRAHN

The data transmission mechanism for the CRHAN is a time-division multiplexing (TDMA) scheme for the data transmission as in [8]. Time is divided into slots of fixed durations. In each time slot, a secondary node is scheduled to send packets. The TDMA scheme is designed with consideration of the interference model to avoid interference as in [8]. We consider two typical routing schemes for CRAHN, one is the routing scheme as defined in [8] that is proved to achieve the best delay performance; the other is a typical routing scheme (and scheduling policy) with queuing and buffering proposed by ourselves, and it is proved of general value and achieving the best throughput capacity. *Routing Scheme 1*: every secondary node in CRAHN will transmit or relay its packets in the time slot allocated to it. The queue length of each nodes are known [12] to be $\Theta(1)$, which is independent with n, and secondary nodes will transmit with full throughput capacity $\Theta(W/\sqrt{n \log n})$. However, collision will occur when secondary node transmissions meet the active primary transmissions, and thus the corresponding secondary data will be lost; *Routing Scheme 2:* when the time slot is allocated to a secondary node, the secondary node determines whether to transmit its traffic or not based on the sensing information of PU activity. When the PU is active, the secondary node will store its traffic in its queue; otherwise it will transmit. As shown in the next analysis, the maximum per-node throughput capacity can be achieved. However, the end-to-end delay is largely increased, and the queue length of each node is dependent to both primary nodes and secondary nodes density.

D. Definitions of Throughput Capacity of CRAHN

We extend the definition of [8] for the per-node throughput to CRAHN, denoted by $\lambda(n, m)$. Due to the impact of PU activity, the per-node throughput capacity should be observed as the time-average of a long time duration: let $\lambda_i(n, m)$ be the total number of bits delivered end-to-end for source-destination pair i up to time t, then the throughput

$\lambda(n, m)$ of the CRAHN is given by

$$\lambda(\mathrm{n}, \mathrm{m}) = \liminf_{t\to\infty} \frac{1}{n} \sum_{i=1}^{n} \frac{\lambda_i(\mathrm{n}, \mathrm{m})}{t}. \tag{5.1}$$

A per-node throughput of λ is said to be feasible if there exists n_0 such that for any $n \geqslant n_0$, there exists a placement rule of PUs, a specific transmission range, PU activity distribution, and a spatial and temporal scheduling scheme for transmission by allowing buffering at intermediate nodes (if necessary) in CRAHN with the property that each bit transmitted by a source is received at its destination with probability at least $1 - \epsilon$, and

$$\lim_{t\to\infty} \Pr\left(\frac{1}{n} \sum_{i=1}^{n} \frac{\lambda_i(\mathrm{n}, \mathrm{m})}{t} \geq \lambda, \forall i \right) = 1. \tag{5.2}$$

E. Definition of Average Packet Delay of CRAHN

The end-to-end delay of a packet is the time it takes to reach the destination after leaving the source. Usually, network delay contains the followings parts [94]: *propagation delay*, *processing delay*, *queuing delay* and *transmission delay*. In [10] [34], multi-hopping delay as the major contributor in end-to-end delay is calculated, and the queuing delay of source node is of constant order $\Theta(1)$, which scales shower than it, is ignored. The multi-hopping delay counts the number of hops passed by the S-D routing path, which can be referred as *processing and propagation delays*.

Unlike classic ad hoc networks, CRAHN may not always transmit due to PU activity: using the Routing Scheme 2, the secondary nodes need to buffer their traffic in their queues when primary nodes are active, so there is no secondary traffic in transmission. As a result, *queuing and transmission delays* occur. Consequently, the end-to-end delay by using Routing Scheme 2 contains three parts: *multi-hopping delay*, *queuing and transmission delays*. The average packet delay of a network is obtained by averaging over all transmitted packets in the network.

5.2 Capacity and Delay of CRAHN under Regular Dense Primary Network

First we analyze the capacity and delay of CRAHN under a typical primary network topology, namely, *regular dense primary network.* The definition is as follows: (1) the m primary transmitters are regularly placed in the unit area square with equal transmission range $r(m)$; (2) all the primary receivers are within its own primary transmitter's transmission range; (3) the density of primary nodes are as dense as possible, such that all of their transmission range are nearly contingent with each other in a "compact" form as illustrated in Fig. 1. We call this primary network as *regular dense primary network*, where nearly all the secondary nodes are covered within primary nodes' transmission ranges. Obviously, this primary network scenario achieves the worst capacity and delay performance in CRAHNs.

In Fig.1, we first divide the unit area square into m *cells* of size $a(m) = 1/m$ where each primary transmitter is located at the center of each cell. Then we divide the unit area square into *micro cells* of size $a(n) = c \log n/n$, where $c \geq 1$. The following lemma guarantees that each micro cell is occupied by at least one secondary node with high probability (w.h.p.). It can be proved by using the property of P.P.P. and the well-known results (for example, see [95], chapter 3). Due to space limitation, we do not repeat the proof here.

Lemma 1: *Divide the network into micro cells of size* $\sqrt{c \log n/n} \times \sqrt{c \log n/n}$*, with* $c \geq 1$*, every micro cell is occupied by at least one secondary node w.h.p.*

For secondary nodes communication, we can set $r(n) = \sqrt{5a(n)}$ such that any secondary node in one micro cell can transmit to the secondary nodes in its four neighboring micro cells. For each source-destination (S-D) pair, we draw a straight line that connects the S-D pair (*S-D line*), and it intersects some micro cells on its path. The source secondary node chooses some of these micro cells intersect with the *S-D line* to relay the traffic to its destination. The corresponding routing path of the S-D pair is illustrated in Fig. 1.

For regular dense primary network, we consider the case where all the secondary nodes are within the primary nodes' transmission ranges w.h.p. It means that the transmission

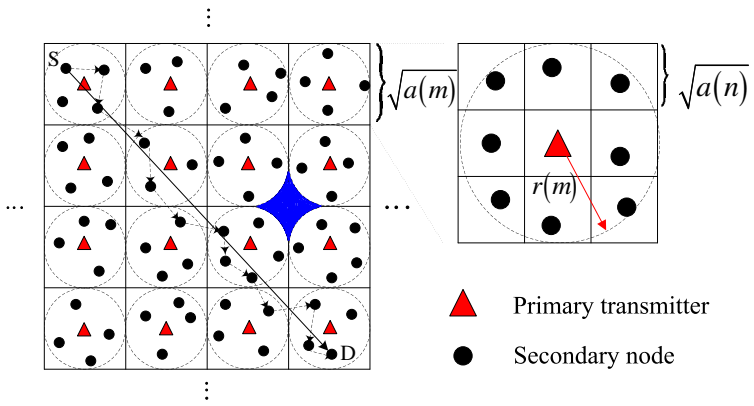

Figure 5.1. CRAHN under regular dense primary network

range satisfies $\sqrt{a(m)}/2 - \sqrt{a(n)}/2 < r(m) \leq \sqrt{a(m)}/2$ to guarantee all the secondary nodes are within these transmission ranges w.h.p. Because the density of primary network is high, the *blue region* in Fig. 1 should not be occupied by secondary nodes w.h.p. By simple calculations, the area of the blue region is $\frac{4-\pi}{4}\frac{1}{m}$. According to Lemma 1, if $\frac{4-\pi}{4}\frac{1}{m} < \frac{c\log n}{n}$, or $m = \omega(n/\log n)$, the probability that the blue region is occupied by at most one secondary node is less than 1 w.h.p. Let us denote this probability by p. So the probability that the source node can find relay nodes that are not covered by primary nodes' transmission region scales to $p^{\Theta(\sqrt{m})}$. When $m \to \infty$, it will approach 0. That means when $m = \omega(n/\log n)$, every secondary S-D pair cannot find a routing path outside the primary nodes' transmission ranges w.h.p. This is the worst case that all the secondary traffic must be affected by the PU activity.

The following lemma is a straightforward conclusion from [8] and [10] that guarantees the interference among secondary nodes is independent of n.

Lemma 2: *The number of micro cells that interfere with any given micro cell is upper bounded by a constant c_3, and is independent of n.*

A consequence of Lemma 2 is that there exists an interference-free schedule where each micro cell becomes active regularly and no micro cell interferes with any other simultaneously transmitting micro cell in CRAHNs.

The following two lemmas follow the proof outline in [10]. We omit them due to space

limitation.

Lemma 3: *For $a(n) = \Omega(\log n/n)$, the number of S-D lines passing through each micro cell are $O(n\sqrt{a(n)})$ w.h.p.*

Lemma 4:*When the primary network is always inactive over time, by using the Routing Scheme 1 with $a(n) \geq \frac{c \log n}{n}$, the throughput and multi-hopping delay of CRAHN become $T(n) = \Theta(W/n\sqrt{a(n)})$ and $D_H(n) = \Theta(1/\sqrt{a(n)})$, i.e., the achievable throughput-delay trade-off is $T(n) = \Theta(\frac{D(n)}{n})$.*

In the next two subsections, we analyze the capacity and delay scaling for CRAHNs by using Routing Schemes 1 and 2.

5.2.1 CRAHN with Routing Scheme 1

Theorem 1: *Under regular dense primary network where $m = \omega(n/\log n)$, using Routing Scheme 1, the per-node throughput capacity of CRAHN is $\Theta(P_{\text{off}}^{\Theta(\sqrt{m})}\frac{W}{\sqrt{n \log n}})$. The corresponding end-to-end delay is $D(n, m) = \Theta(\sqrt{n/\log n})$.*

Proof Since the transmission range of each primary node is $\Theta(\sqrt{a(m)}) = \Theta\left(1/\sqrt{m}\right)$, the number of primary cells passed by a packet from a secondary node through the S-D line i is $\Theta(l_i/\sqrt{a(m)})$, where l_i is the length of S-D line i. Thus, the number of primary cells passed by a packet from a secondary node averaged over all S-D pairs of CRAHNs is $\Theta(\frac{1}{n}\sum_{i=1}^{n} l_i/\sqrt{a(m)})$. Since m and n are independent variables, for large n, the average length of S-D lines is $\Theta\left(\frac{1}{n}\sum_{i=1}^{n} l_i\right) = \Theta(1)$. So the average number of primary cells passed by a packet from a secondary node averaged over all S-D pairs is $\Theta(1/\sqrt{a(m)}) = \Theta(\sqrt{m})$.

From Lemma 4, the maximum traffic rate of a source secondary node on each S-D line is $\Theta(W/n\sqrt{a(n)}) = \Theta(W/\sqrt{n \log n})$. Using Routing Scheme 1, the secondary nodes need to transmit their packets when its time slot is available. The condition for successful packets transmission is the case that all the primary cells along its routing path is inactive, or collision occurs and secondary data are lost. Since the average number of primary cells passed by a packet averaged over all S-D pairs is $\Theta(\sqrt{m})$ and each primary cell's activity is independent of others, the average probability of successful packets transmission from

source to destination for each S-D pair is $P_{off}^{\Theta(\sqrt{m})}$. So the per-node throughput capacity of CRAHN is

$$\lambda(n,m)=\Theta\left(P_{off}^{\Theta(\sqrt{m})}\frac{W}{n\sqrt{a(n)}}\right)=\Theta\left(P_{off}^{\Theta(\sqrt{m})}\frac{W}{\sqrt{n\log n}}\right). \tag{5.3}$$

In this case, there is only multi-hopping delay as in classical ad hoc wireless networks. According to lemma 4, $D(n,m)=D_H(n)=\Theta(\sqrt{n/\log n})$.

From equ.(3), we can see, compared to the classic ad hoc networks, the per-node throughput capacity of CRAHN will drain rapidly at an exponential rate with m when P_{off} is fixed or at a polynomial rate with P_{off} when m is fixed in the condition of no additional queueing and buffering strategy in the regular dense primary scenario. This results reveal the strong dependence of primary network topology and CRAHN performance. It suggests the necessity of novel protocol design for CRAHNs to improve the throughput capacity.

5.2.2 CRAHN with Routing Scheme 2

We consider the throughput capacity and delay scaling for CRAHN by using the Routing Scheme 2, where secondary nodes buffer their traffic in queues when the corresponding primary cell is active; otherwise, they can transmit. We first investigate the throughput capacity and delay trade-off, and the required average queue length in each secondary node. Then, as a direct result, the maximal achievable per-node throughput capacity, end-to-end delay, and queue length are present.

Theorem 2: *Under regular dense primary network where $m=\omega(n/\log n)$, by using Routing Scheme 2, the per-node throughput capacity of CRAHN and the queueing and transmission delays trade-off are given by*

$$\lambda(n,m)=\Theta\left(\frac{WP_{\text{off}}^{\Theta(\sqrt{m})/K}}{\sqrt{n\log n}}\right),\ \text{for}\ \mathrm{D_{QT}(n,m)=O(K)},$$

where $1\le K\le\Theta(\sqrt{m})$. The average queue length for each secondary node is $Q(n,m)=O\left(\frac{m}{\sqrt{n}K}P_{\text{off}}^{\Theta(\sqrt{m})/K}\right)$.

Proof In Fig.1, we denote the length of the S-D line is l_i. So the number of primary cells that the S-D line passes is $N=\Theta(l_i/\sqrt{a(m)})$. Fig. 2 illustrates the process of how to

calculate the throughput of the S-D pair. In Fig. 2.(a), the source secondary traffic passing through the N primary cells to its destination both in time and space. In the *time axis*, $T = \alpha + \beta$, denotes expected active and inactive time of primary nodes. $T_{on} = TP_{on} = \alpha$ represents the expected time that the primary cell is active. Similarly, $T_{off} = TP_{off} = \beta$ represents the expected time that the primary cell is inactive. In the *space axis*, we denote the arrival and departure rates of the secondary traffic along the N primary cells by $\lambda_0, \lambda_1, ..., \lambda_N$. Especially, λ_0 and λ_N are the source traffic and destination traffic rates, respectively. For tractable analysis of secondary traffic, we can neglect the multi-hopping delay as it scales slower than the queuing and transmission delays w.h.p. The relation between them is justified later in this subsection thus validating this assumption. Without the loss of generality, we consider the time that the source secondary node is blocked by the primary cell 1 at time 0 as illustrated in Fig. 2.(a). At first, the source secondary node needs to buffer its traffic during the time interval $[0, T_{on}]$, since the primary cell is active. Then, the source secondary node needs to transmit with the rate $\lambda_1 > \lambda_0$ to clear this queue within the time interval $[T_{off}, T]$. The traffic generated by the source secondary node in the period T can be either forwarded to next intermediate secondary nodes or destination secondary node in other primary cells. Note that, if the queue of secondary nodes in the primary cell 1 is not cleared within the time period T and suppose the remaining queue length is ΔQ, in the next on-and-off time circle T of the primary cell 1, the remaining queue length will be $2\Delta Q$. As this process is repeated after NT, the remaining queue length in the source secondary node is $NT\Delta Q$. As $N \to \infty$, the queue length will approach infinity, which makes its queue overflow. This point is illustrated in Fig. 2. (b). At this time, we have

$$\frac{\lambda_0 T_{on}}{\lambda_1 - \lambda_0} \leq T_{off}. \tag{5.4}$$

By rewriting this equation, we have $\lambda_0 \leq \lambda_1 P_{off}$. Since the maximum achievable throughput capacity of λ_1 is $\Theta(W/\sqrt{n \log n})$, so according to the definition in equ. (2), the feasible throughput capacity of the source secondary node that guarantees stable transmission over time is upper bounded by $O\left(P_{off} W/\sqrt{n \log n}\right)$.

Note that, if all the left primary cells along the S-D line during the time interval $[T_{off}, T]$ are inactive, the source secondary node can transmit its data smoothly to its destination with queuing and transmission delays only as T. The source secondary node can achieve the maximum throughput capacity $\Theta\left(P_{off}W/\sqrt{n\log n}\right)$. However, this scenario is the ideal case where the primary cells' on-and-off activity is totally the same as the others which violate the independence assumption of the primary cells. The general case is that, when secondary source node intends to transmit its data in $[T_{off}, T]$, the location of other primary cells' which are active at the moment, as in Fig. 2.(a), denoted by $t_1, t_2, t_3, ..., t_{N-1}$, is uniformly distributed and independent of each other. As denoted by the blue axis in Fig. 2. (a), the secondary traffic can only pass through during the intersection of inactive time intervals, denoted by t_{off}, along the concatenated primary cells. Or, the secondary traffic will be blocked in the queue. One observation is that, the more primary cells the source secondary traffic passing through, the less the intersection of the inactive time interval of t_{off}. Similar to eq.(5.4), we have $\lambda_0 T_{on} \leq (\lambda_1 - \lambda_0)t_{off}$. Thus, the source secondary traffic rate will be reduced. However, the corresponding end-to-end queuing and transmission delays will be small.

According to the above analysis, in the next, we investigate how the source secondary traffic passing through K successive primary cells during one on-and-off time cycle T, where $K \in \{1, 2, ..., N\}$, and repeats this process until the secondary traffic reaches its destination. This scheduling policy is indicated in the following discussion to be suitable for all the possible on-and-off distribution of PU activity. Thus this scheduling policy is of general value, and the throughput capacity (or called "traffic rate" in the queuing analysis) and delay tradeoff in CRAHN can be obtained for general conditions.

5.2.2.1 Capacity, Delay and Queue Length with $K = 1$

We first consider the case where $K = 1$. Fig. 2.(c) illustrate how the S-D secondary traffic transmits across the primary cells and achieves the source throughput λ_0. In Fig.2.(c).(1), the source secondary traffic is blocked in the primary cell 1 during the time $[0, T_{on}]$, since

the primary nodes are active. So the source secondary node in primary cell 1 should have a queue of length $\lambda_0 T_{on}$ to store its traffic. During the time $[T_{on}, T]$, the secondary node transmits with the traffic rate $\lambda_1 > \lambda_0$ to clear its queue and transmits all the traffic to next secondary nodes in the primary cell 2. The secondary traffic transmitted from primary cell 1 to 2 is $\lambda_1 T_{off} = \lambda_0 \Big/ P_{off} \cdot T_{off} = \lambda_0 T$. Next, in Fig.2.(c).(3), during the time $[T, T + T_{on}]$, the source secondary node is injecting new traffic into primary cell 1 with queue length $\lambda_0 T_{on}$, and the secondary traffic generated during the time $[0, T]$ with the size $\lambda_0 T$ is transmitted to primary cell 2 during the inactive time of the primary cell 2 where $\lambda_0 T = \lambda_2 \mathrm{T}_{off}$. Since $\lambda_0 = \Theta(P_{off} W/\sqrt{n \log n})$, we have $\lambda_2 = \Theta(W/\sqrt{n \log n})$. Next, in Fig.2.(c).(4), during the time $[T + T_{on}, 2T]$, the secondary traffic with queue length $\lambda_0 T$ buffered in primary cell 1 will be transmitted to primary cell 2 with traffic rate $\Theta(W/\sqrt{n \log n})$; similarly, the secondary traffic buffered in primary cell 2 with queue length $\lambda_0 T$ will be transmitted to primary cell 3 with the traffic rate $\Theta(W/\sqrt{n \log n})$. By repeating this process for time of NT, the source secondary traffic will reach its destination nodes.

The stable phase of the queues on the S-D line is shown in Fig.2.(d). The traffic rate of $\lambda_1, ..., \lambda_N$ equal to $\Theta(W/\sqrt{n \log n})$ within primary inactive time slots and equal to zero within primary active time slots. The feasible source traffic rate in the stable phase is $\lambda_0 = \Theta(P_{off} W/\sqrt{n \log n})$. The average per-node throughput capacity of CRAHN is the average maximal traffic rate of all S-D pairs transmitted from source to destination nodes. So, according to the fluid model of traffic [10], the per-node throughput capacity can be referred as the feasible traffic rate generated from source node that guarantees the stable traffic on its routing path, that is

$$\lambda(n, m) = \Theta\left(\frac{W P_{off}}{\sqrt{n \log n}}\right). \tag{5.5}$$

Next we compute the average packet queuing and transmission delays. As we already analyzed, a packet needs the time $NT = \Theta(l_i/\sqrt{a(m)} \cdot T)$ to reach its destination. The average queuing and transmission delays over all S-D pairs in CRAHN is $NT = \Theta(\frac{1}{n}\sum_{i=1}^{n} l_i/\sqrt{a(m)} \cdot T)$. For large n, the average distance between S-D pairs is $\frac{1}{n}\sum_{i=1}^{n} l_i =$

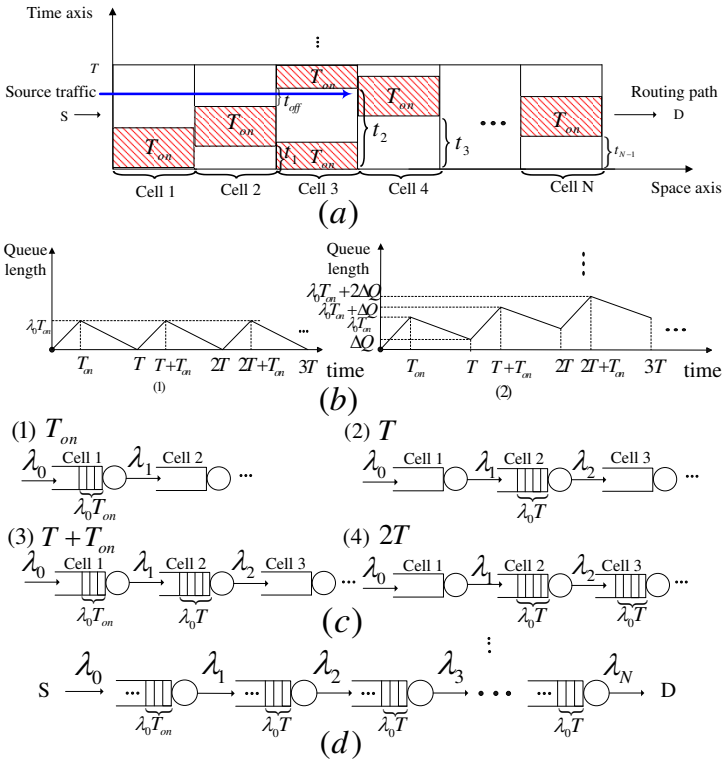

Figure 5.2. Secondary traffic passing through S-D line using routing scheme 2

$\Theta(1)$.

Note that, the proposed scheduling strategy are applicable for the worst distribution of primary on-and-off activity for CRAHN, where $P_{on} \gg P_{off}$ and $P_{on} \to 1$. In this case, the secondary traffic can only be buffered during every T, and then be transmitted to the secondary nodes in the next primary cell, This is the exact buffering and scheduling policy we considered above.

For general P_{off} and P_{on}, i.e., $P_{on} \ll P_{off}$, if $P_{on} \to 0$, the intersection of inactive time along any K ($2 \le K \le N$) successive primary cells for the limiting case is $t_{off} \to T$. Thus the source secondary traffic has chance to pass through more primary cells during T, and the queuing and transmission delays is reduced. So the average queuing and transmission delays is consequently upper bounded by:

$$D_{QT}(n,m) = O\left(\frac{1}{\sqrt{a(m)}} \cdot T\right) = O\left(\sqrt{m}T\right) = O\left(\sqrt{m}\right). \tag{5.6}$$

Note that, the multi-hopping delay for CRAHN is $D_{\mathrm{H}}(n) = \Theta(\sqrt{n/\log n})$ according to lemma 4. From the definition of the regular dense primary network, to guarantee all the secondary nodes to be covered by primary cells w.h.p., we have $m = \omega(n/\log n)$. so $\sqrt{m} = \omega(\sqrt{n/\log n})$. Compared with eq.(6), this means that the queuing and transmission delay $D_{QT}(n,m)$ scales faster than the multi-hopping delay $D_{\mathrm{H}}(n)$, which is the major contributing factor to the end-to-end delay, and the delay caused by multi-hopping can be omitted. Thus, the end-to-end delay is computed from

$$D(n,m) = O(\sqrt{m}). \tag{5.7}$$

Next we compute the queue length needed for the buffering of secondary traffic. As shown in Fig.2.(d), in the stable phase, the secondary nodes along the path of N primary cells need queues to store their traffic, where the length of the queue required in the primary cell 1 is $\lambda_0 T_{on}$. In the remaining $N-1$ primary cells, the required queue length in each primary cell is $\lambda_0 T$. In general, P_{off} or P_{on}, the total queue length along the S-D line is upper bounded by $O(\lambda_0 T_{on} + (N-1)\lambda_0 T)$. Since there are $\Theta(l_i/\sqrt{a(n)})$ secondary nodes along the S-D line, the average primary cells and secondary nodes over all S-D pairs of CRAHN are $\Theta(\frac{1}{n}\sum_{i=1}^{n} l_i/\sqrt{a(m)}) = \Theta(\sqrt{m})$ and $\Theta(\frac{1}{n}\sum_{i=1}^{n} l_i/\sqrt{a(n)}) = \Theta(\sqrt{n})$. So the average queue length of each secondary node per S-D pair i is

$$\begin{aligned} Q_{\text{S-D}}(n,m) &= \frac{O(\lambda_0 T_{on} + (N-1)\lambda_0 T)}{\Theta(\sqrt{n})} = \frac{O(N\lambda_0 T + \lambda_0 T_{off})}{\Theta(\sqrt{n})} \\ &= \frac{O(\sqrt{m}\lambda_0 T)}{\Theta(\sqrt{n})} + \frac{O(\lambda_0 T_{off})}{\Theta(\sqrt{n})} \\ &= O\left(\sqrt{\frac{m}{n}}\right)\lambda_0 T + O\left(\frac{1}{\sqrt{n}}\right)\lambda_0 T_{off} \\ &= O\left(\sqrt{\frac{m}{n}}\lambda_0 T\right) = O\left(\sqrt{\frac{m}{n}}\frac{WP_{off}T}{\sqrt{n\log n}}\right), \end{aligned} \tag{5.8}$$

where W and T are constants. So the average queue length of each secondary node along every S-D pair can be further simplified as $O(\sqrt{\frac{m}{n}}\frac{P_{off}}{\sqrt{n\log n}})$. From Lemma 3, we know the number of S-D lines passing through each micro cell is $\mathrm{O}(n\sqrt{a(n)}) = O\left(\sqrt{n\log n}\right)$. So for each secondary node, the total queue length required to support all the S-D lines passing

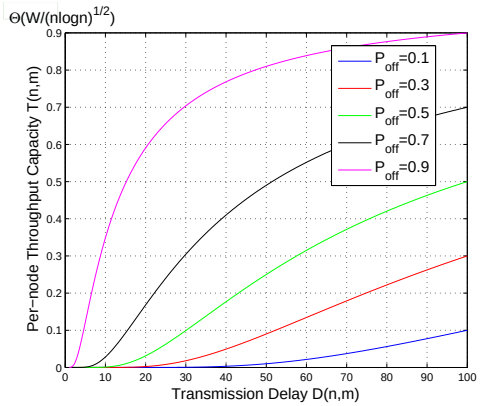

Figure 5.3. Capacity and delay tradeoff

through it, is given by

$$\begin{aligned} Q(n,m) &= Q_{\text{S-D}}(n,m) * O\left(\sqrt{n\log n}\right) \\ &= O\left(\sqrt{\frac{m}{n}}\frac{P_{off}}{\sqrt{n\log n}}\right) * O\left(\sqrt{n\log n}\right) \\ &= O\left(\sqrt{\frac{m}{n}}P_{off}\right). \end{aligned} \tag{5.9}$$

5.2.2.2 *Capacity, Delay and Queue Length with $K \geq 1$*

First we consider the case where $K = 2$. We need to estimate the secondary traffic rate after passing through the first two primary cells where the beginning of active time of primary cell 2 is the random variable t_1 with uniform distribution $t_1 \sim U(0,T)$. To estimate the intersection time t_{off} of primary cell 1 and 2 for all the S-D pairs, according to central limit theorem [95], we have $\frac{1}{n}\sum_i^n t_{off}^i \to E[t_{off}], n \to \infty$. By taking the expectation $E\left[t_{off}\right] = T_{off}^2/T$, we have $\lambda_0 \leq E\left[t_{off}\right]\lambda_1/T = P_{off}^2\lambda_1$. Since we need to obtain the maximum source throughput capacity λ_0. We can set $\lambda_0 = P_{off}^2\lambda_1$. In this case, the transmitted traffic which is buffered in the queue of the intermediate secondary node in the primary cell 2 is $\lambda_0 T$. So for all the S-D pairs, the estimated maximum achievable source traffic rate is $\lambda_0 = E\left[t_{off}\right]\Theta(W/\sqrt{n\log n})/T = \Theta(WP_{off}^2/\sqrt{n\log n})$.

For general K ($K \geq 2$), similarly as in $K = 1$, the source secondary traffic of size $\lambda_0 T$ will be transmitted and buffered at the secondary nodes in the Kth primary cells within the time cycle T. Then the traffic will be transmitted and buffered at the next K primary cells

within the next time cycle T. We repeat this process until the traffic reaches the destination node. By using a simple mathematical induction, we can get $E\left[t_{off}\right] = P_{off}^{K-1}T_{off}/T$, $\lambda_0 \leq E\left[t_{off}\right]\lambda_K/T = P_{off}^{K}\lambda_K$. In this case, the per-node throughput capacity becomes:

$$\lambda(n,m) = E[t_{off}]\Theta(\frac{W}{\sqrt{n\log n}})/T = \Theta(\frac{WP_{off}^{K}}{\sqrt{n\log n}}). \tag{5.10}$$

Because the secondary traffic passes K primary cells with every T time, the corresponding queuing and transmission delays is

$$D_{QT}(n,m) = O\left(\sqrt{m}T\right)/K = O\left(\frac{\sqrt{m}}{K}\right). \tag{5.11}$$

Similar as $K = 1$, in this case, there are only $\Theta(\sqrt{n}/K)$ secondary nodes to buffer secondary traffic along the S-D line. The average total queue length required for each secondary node to support all the S-D lines passing through it, is given by

$$\begin{aligned} Q(n,m) &= Q_{S\text{–}D}(n,m) * O\left(\sqrt{n\log n}\right) \\ &= \frac{\lambda_0 T_{on}+(N-1)\lambda_0 T}{\Theta(\sqrt{n}/K)} * O\left(\sqrt{n\log n}\right) \\ &= O\left(\sqrt{\frac{m}{n}}KP_{off}^{K}\right). \end{aligned} \tag{5.12}$$

Let $K = \Theta(\sqrt{m})/K$, we obtain the results in Theorem 2.

From Theorem 2, we can draw the capacity and delay trade-off curve. As an example, in Fig. 3 we show the results for $m = 100$ and different values of P_{off}.

Substituting $K = 1$ in Theorem 2, we can obtain the Theorem 1 as a special case. Note that now the multi-hopping delay $D_{\mathrm{H}}(n) = \Theta(\sqrt{n/\log n})$ scales faster than $D_{QT}(n,m) = O(1)$, which is the major contributing factor to the end-to-end delay.

The following theorem claims the maximum achievable throughput capacity of CRAH-N using the Routing Scheme 2.

Theorem 3: *Under the regular dense primary network where $m = \omega(n/\log n)$, using the Routing Scheme 2, the maximum achievable per-node throughput capacity for CRAHN is $\lambda(n,m) = \Theta\left(\frac{WP_{\mathrm{off}}}{\sqrt{n\log n}}\right)$, each secondary node needs a queue length of order $Q(n,m) = O\left(\sqrt{\frac{m}{n}}P_{\mathrm{off}}\right)$ to buffer traffic. The corresponding end-to-end delay is $D(n,m) = O(\sqrt{m})$.*

Proof Substituting $K = \Theta(\sqrt{m})$ in Theorem 2, we can directly derive this theorem. Also, as a special case, the detailed steps of the derivation can be constructed from the proof of Theorem 2 in Section III.*B.1)*.

From Theorem 3, we can determine that without packet losses and having enough queue length to combat the packet loss due to the PU activity, the maximum achievable per-node throughput capacity can be greatly improved from an exponentially draining factor $P_{off}^{\Theta(\sqrt{m})}$ to a linear factor P_{off} when compared with results by using the Routing Scheme 1. To achieve this throughput capacity improvement, It needs additional queue length in the order of $O\left(\sqrt{\frac{m}{n}}P_{off}\right)$. Moreover, the end-to-end queuing and transmission delays are largely increased to $O(\sqrt{m})$. Since the delay perceived by a packet is more QoS relevant than the total network throughput, we need to use the results of Theorem 2 as a guidance to get the best throughput and delay trade-off. For $Q(n, m)$, when increasing the density m of primary nodes, the queue length of secondary nodes will increase in the order of $\sqrt{m}$. Otherwise when increasing the density n of secondary nodes, the queue length of the secondary nodes will decrease in the order of $\sqrt{n}$. This result coincides with the intuitive observation. From Theorem 3, we can also determine that when $P_{off} = 0$, that means all the primary nodes are active, no secondary nodes can transmit at all. Thus, the capacity of CRAHN is zero and the queue length of each secondary node is zero. When $P_{off} = 1$, the primary nodes are always inactive. In this case, CRAHN behaves the same as a traditional ad hoc wireless network. These special cases verify our results in Theorem 3.

5.3 Capacity and Delay of CRAHN under Regular Sparse Primary Network

In this section, we consider another typical scenario similar to the regular dense primary network, where the primary nodes' density is low and transmission range is small than the dense primary networks. That is $m = o(n/\log n)$; the transmission range $0 \leq r(m) \leq$

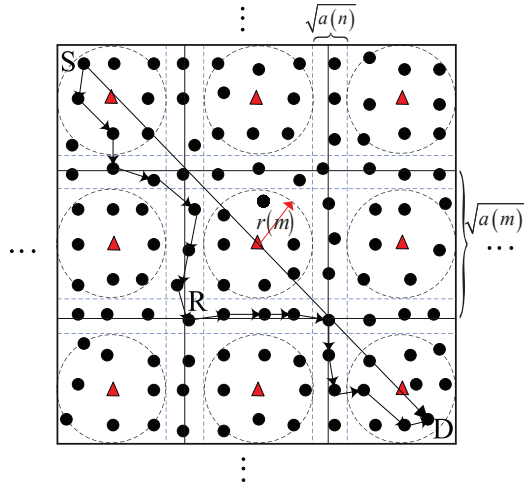

Figure 5.4. CRAHN under regular sparse primary network

$\sqrt{a(m)}/2 - \sqrt{a(n)}/2$ guarantees there are always secondary nodes outside of the transmission ranges of primary cells, and finding a routing path outside of primary transmission range (besides source and destination cells) is proved to be always feasible (in lemma 4). In Fig. 4, the typical routing path S-R-D that avoid primary nodes' transmission range is named as *detour routing*. This network is defined as *regular sparse primary network*. The following lemma proves that the detour routing scheme has the same order of routes traverse any micro cell as the S-D line routing schemes in lemma 3.

lemma 5*: The number of routes in detour routing scheme that traverse any micro cell are* $O(n\sqrt{a(n)})$.

Proof From lemma 3 we know that the number of S-D lines traversing a single micro cell are $O(n\sqrt{a(n)})$. Now, consider the number of routes of using S-R-D routing scheme that passed through the uncovered secondary micro cells. Since it is enough to fill all the S-R-D line on all the uncovered region of secondary micro cells, In the worst case where primary nodes have the largest transmission range as in Fig. 3, the area of these uncovered region is $1 - \mathrm{m}\pi\left(\left(\sqrt{a(m)} - \sqrt{a(n)}\right)/2\right)^2 = \Omega(1)$, when $m = O(n/\log n)$. So, the number of route of the detour routing scheme that traverse any cell is still $\frac{O(n\sqrt{a(n)})}{1-\mathrm{m}\pi\left(\left(\sqrt{a(m)}-\sqrt{a(n)}\right)/2\right)^2} = O(n\sqrt{a(n)})$.

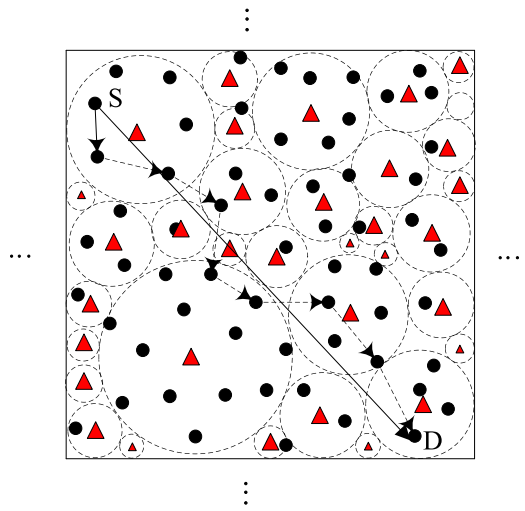

Figure 5.5. CRAHN under random dense primary network

Theorem 4: *Under regular sparse primary network where* $m = o(n/\log n)$ *and the primary nodes' transmission range* $0 \leq r(m) \leq \sqrt{a(m)}/2 - \sqrt{a(n)}/2$, *the maximum achievable per-node throughput capacity of CRAHN* $\lambda(n,m)$ *is bounded between* $\Omega(\frac{WP_{\text{off}}}{\sqrt{n\log n}})$ *and* $O(\frac{W}{\sqrt{n\log n}})$. *The corresponding end-to-end delay is mainly the multi-hopping delay,* $D(n,m) = \Theta(\sqrt{n/\log n})$.

Proof Let us denote the length of the straight line that connects the S-D pair is l_i as in Fig. 4. For any S-R-D pair, the length of its detour routing path, as shown in Fig. 4, is at most $\sqrt{2}l_i$. As we known [10], the constant difference of the length of routing path does not change the delay scaling order. Moreover, from lemma 5, we know the detour routing scheme does not change the route numbers in each each micro cells in the uncovered region of primary nodes. Ignoring the impact of PU activity on its covered secondary nodes, according to the TDMA transmission scheme, only one out of $O(n\sqrt{a(n)})$ S-D pairs can transmit with no interference to other S-D pairs. So the throughout capacity is upper bounded by (and the upper bound is achievable) $1/O(n\sqrt{a(n)})$, which is the same as the results in classic wireless networks [10].

Now we consider the scalability of CRAHN with the all possible locations of nodes:

source and/or destination secondary nodes are covered or uncovered in the PUs' transmission ranges. Let X be the Poisson random variable of secondary nodes density with mean n and standard deviation $\sqrt{n}$ in the square region. The mean and standard number of secondary nodes in a primary node's transmission range is $n\pi r^2(m)$ and $\sqrt{n\pi r^2(m)}$. Recall the Chebyshev's inequality [95]:

- For any $t > 1$, $\Pr(|X - E[X]| \geq t \cdot \sigma[X]) \leq \frac{1}{t^2}$.

Accordingly, for any $\varepsilon > 0$ and fixed $r(m)$, we can obtain

$$\begin{aligned} &\Pr\left(|X - n\pi r^2(m)| \geq (\varepsilon \sqrt{n\pi r^2(m)}) \sqrt{n\pi r^2(m)}\right) \\ &\qquad \leq \frac{1}{(\varepsilon \sqrt{n\pi r^2(m)})^2} = \frac{1}{\varepsilon^2 \pi r^2(m) \cdot n}. \end{aligned} \tag{5.13}$$

As $n \to \infty$, we can make this quantity arbitrarily small. So the number of secondary nodes in any primary node's transmission range is $n\pi r^2(m)$. Similarly, when $n \to \infty$, we can prove the number of secondary nodes in a primary cell (square region) is n/m. So the probability for a secondary node in any primary node's transmission range is $P_{in} = \pi r^2(m)/\frac{1}{m}$; the probability for a secondary node out of any primary node's transmission range is $P_{out} = 1 - \pi r^2(m)/\frac{1}{m}$. Note that, to calculate the throughput capacity of CRAHN, we choose the source and destination nodes randomly. So there are four possible cases of choosing S-D pairs:

- Source and destination nodes are both in primary nodes' transmission range, then the probability is $P_{in}P_{in}$
- Source node is in a primary nodes' transmission range, destination node is out of primary nodes' transmission range, then the probability is $P_{in}P_{out}$
- Source node is out of primary nodes' transmission range, destination node is in primary nodes' transmission range, then the probability is $P_{out}P_{in}$
- Source and destination nodes are both out of primary nodes' transmission range, then the probability is $P_{out}P_{out}$.

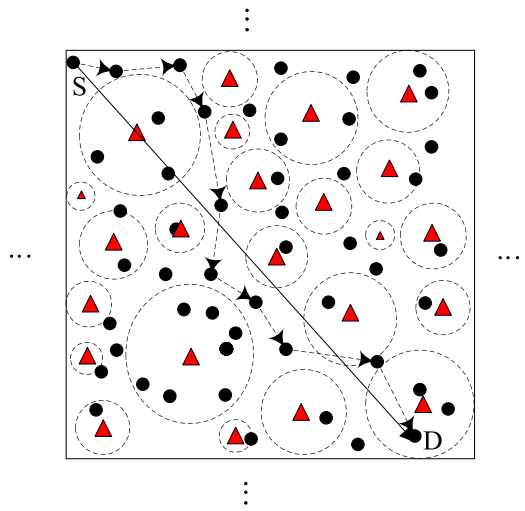

Figure 5.6. CRAHN under random sparse primary network

When Routing Scheme 2 is used, then the case (1) is a special case in the proof of Theorem 3, where $m = 2$. So the capacity is $\Theta(WP_{off}/\sqrt{n\log n})$, and the queuing and transmission delays is $2T$. Case (2) and (3) are special cases in the proof of Theorem 3, where $m = 1$. The capacity is $\Theta(WP_{off}/\sqrt{n\log n})$, and the queuing and transmission delays is T. Case (4) is the same as in classical ad hoc network scenarios, so the capacity is $\Theta(W/\sqrt{n\log n})$. As a result, the average per-node throughput capacity of CRAHN is

$$\begin{aligned}\lambda(n,m) = {} & P_{in}^2\Theta\left(\frac{WP_{off}}{\sqrt{n\log n}}\right) + P_{out}^2\Theta\left(\frac{W}{\sqrt{n\log n}}\right) \\ & +2P_{in}P_{out}\Theta\left(\frac{WP_{off}}{\sqrt{n\log n}}\right).\end{aligned} \tag{5.14}$$

From eq.(5.14), we can determine that the throughput capacity of CRAHN is tightly related to the value of primary nodes' density and the transmission range. When $r(m) = 0$ or $m \to 0$, CRAHN achieves the maximum capacity $\Theta(W/\sqrt{n\log n})$; when $r(m) \to \sqrt{a(m)}/2 - \sqrt{a(n)}/2$ or $m = \Omega(n/\log n)$, CRAHN achieves the minimum capacity $\Theta(WP_{off}/\sqrt{n\log n})$.

For the required queue, it is only demanded by the secondary nodes which are covered within source and destination primary cells. They are constant values, and their length will approach zero as $n \to \infty$. To compute the end-to-end delay, we have the multi-hopping delay scaled as $\Theta(\sqrt{n/\log n})$, which scales faster than the queuing and transmission delays $O(2T)$. So the end-to-end delay is:

$$D(n,m) = \Theta\left(\sqrt{\frac{n}{\log n}}\right). \tag{5.15}$$

Compare to Routing Scheme 2, the main difference of Routing Scheme 1 is the difference of lower bound capacity $\Omega(WP_{off}^2/\sqrt{n\log n})$. Note that the upper bound capacity and end-to-end delay remain the same for both schemes. Thus, we complete the proof, and the maximum achievable throughput capacity is achieved by using Routing Scheme 2.

From Theorem 4, we can observe significant improvement of network throughput capacity and delay by the detour routing scheme, since the impact of PU activity on the secondary routing path is largely reduced.

5.4 Capacity and Delay of CRAHN under Random Primary Network

In the last sections III and IV, we considered the regular primary network where secondary nodes are homogeneous nodes with equal transmission ranges placed regularly in the network. Generally, the primary nodes can be randomly distributed in a network region, and the location and transmission range of each primary node might be different from each other. We call these networks as *random primary networks*. The definition are similar as regular primary networks, where the density and transmission range of them have the same scalability. However, the location of primary nodes are randomly distributed now. The corresponding *dense* and *sparse* primary networks for CRAHN are illustrated in Figs. 5 and 6.

Theorem 5: *The throughput capacity and delay of CRAHNs under random (dense or sparse) primary networks scale the same order as in regular (dense or sparse) primary networks, respectively.*

Proof As a typical illustration, we consider there are m primary nodes randomly distributed in the unit square region. The transmission range of a primary node follows a uniform distribution $r(m) \sim U\left(r/\sqrt{m}, R/\sqrt{m}\right)$, where r and R are the lower and upper bound constants. For the random dense primary network, the primary nodes are randomly located

and are grouped together in a dense form as shown in Fig. 5. The value of r and R need to satisfy the following two conditions:

$$0 \le r \le \frac{1}{2} \le R \le 1; \quad r^2 + R^2 + Rr \le 1. \tag{5.16}$$

to guarantee a valid distribution of transmission ranges in this region. For the S-D line with length l_i shown in Fig. 5, the number of primary cells passed through is given by

$$\frac{1}{R/\sqrt{m} - r/\sqrt{m}} \int_{\frac{r}{\sqrt{m}}}^{\frac{R}{\sqrt{m}}} \frac{l_i}{2r\,(m)} dr\,(m) = \left(\frac{l_i}{2} \frac{1}{R-r} \ln \frac{R}{r}\right) \sqrt{m}. \tag{5.17}$$

Thus, the number of primary cells passed by a packet averaged over all S-D pairs in CRAHN is $\Theta(\frac{1}{n} \sum_{i=1}^{n} \frac{l_i}{2} \frac{1}{R-r} \ln \frac{R}{r} \cdot \sqrt{m}) = \Theta(\frac{1}{2(R-r)} \ln \frac{R}{r} \cdot \sqrt{m}) = \Theta(\sqrt{m})$. Since only the number of primary cells along S-D lines impacts the scaling performance (both throughput capacity and delay) of CRAHN and the size of the transmission range does not impact the scalability order, the scaling laws of CRAHN in random dense primary network are similar to the regular dense primary network cases. There is only a constant difference $\frac{1}{2(R-r)} \ln \frac{R}{r}$ which does not change the scalability. Due to similar reason, as in Fig. 5, we can prove that the scaling laws of CRAHN under random sparse primary network are the same as the regular sparse primary network case.

CHAPTER 6

CONCLUSION

In this dissertation, we have studied power control and capacity analysis for cognitive radio networks. The main contributions are summarized as follows.

6.1 Power Control in Cognitive Radio Networks

The main results and conclusions of the proposed power-control problems are summarized as follows.

6.1.1 Power Control in Wideband Cognitive Radio Networks

In this work, a joint power- and rate-control problem is considered for individual SU in wideband CRNs through a game theoretical approach. Proper utility function that guarantees the fairness and energy efficiency among SUs has been adopted. The formulated joint power- and rate-control problem that satisfies both the interference power constraints and QoS requirements of each SU is difficult to solve through classic optimization theory. To characterize the interference, the concept of pricing is introduced from game theory into the utility functions in the analysis. Thus, the proposed problem is a non-cooperative joint power- and rate-control game with interference power pricing. The existence, uniqueness, and Pareto efficiency of Nash equilibrium for this game are presented. By using the analytical results, a distributive joint power- and rate-control algorithm is proposed to finds optimal rate and power for each SU. Moreover, the feasible pricing problem for the CRN is discussed. The numerical results show that our proposed algorithm can provide larger transmission rates for SUs with less transmit powers when compared to conventional power-control problems in CRNs that proves the *joint power and rate control* is an effective approach to improve utility and spectrum utilization of CR networks. Moreover, introducing the pricing into the joint power- and rate-control problem can largely improve the network utility.

6.1.2 Power Control in Fading Multi-Channel Cognitive Radio Networks

In this research, the distributed multi-channel power-control problem is investigated for CR networks over fading channels with channel outage constraints. The fundamental performance traits of multiple CR user power control in the presence of multiple PUs in fading channels is presented for the first time. The properties of N.E. for our game are analyzed, and a distributed power-control algorithm is proposed to solve the game, which converges to the N.E. Finally, the performance of our algorithms are investigated through simulation results, which demonstrates the results shown in theory.

6.1.3 Learning-based Power-Control Game Cognitive Radio Networks

In this work, a robust power-control algorithm with low implementation complexity is designed for *competitive* and *autonomous* CR networks for the first time. The general feature of "asymptotically stationary in the average sense" property is modeled in wireless environments. This formulated problem is an incomplete-information repeated game with learning automata, and the interesting property of the asymptotically equivalence of the classic complete-information case and the incomplete-information power-control game proposed in this work are established in this study.

Unlike traditional approaches, the game-theoretical problem is formulated as a strictly diagonal concave game, which proves the uniqueness of the power-control game with nice mathematical structure and interpretation. To solve this mixed-strategy repeated power-control game, Bush-Mosteller reinforcement scheme is proposed, in which the reinforcement procedure that uses the Lagrange multipliers and an appropriate regularization is shown to be the optimal response within the given constraints for each CR user. Using the stochastic optimization techniques, the *convergence* of the game to the Nash equilibrium is analyzed with the *rate of the learning* to be $o(n^{-1/3})$. This proposed algorithm is very adaptable and can be applicable to any hostile radio-access environments and uncoordinated CR users for the practical applications of real-world CR networks.

6.2 Capacity Analysis in Cognitive Radio Networks

The main results and contributions of the fundamental throughput capacity scaling for large-scale wireless ad hoc networks are summarized as follows:

- Under the regular dense primary network, using TDMA-based routing scheme as in [8] where PU activity is not considered, the per-node throughput capacity of CRAHN is only $\Theta\left(P_{off}^{\Theta(\sqrt{m})}\frac{W}{\sqrt{n\log n}}\right)$. In this case, the throughput capacity drains exponentially with the increasing density m of primary nodes. The corresponding end-to-end delay is only for multi-hopping as in [10].

- When secondary nodes consider PU activity, and use additional queue to buffer traffic when primary nodes are active, otherwise they can transmit, then the maximum achievable throughput capacity is improved to $\Theta(\frac{WP_{off}}{\sqrt{n\log n}})$ which scales down by a factor P_{off} $(0 \leq P_{off} \leq 1)$ compared to classical wireless ad hoc networks. Moreover, to achieve this throughput capacity improvement, the end-to-end delay capturing queuing and transmission delays, is increased to $O\left(\sqrt{m}\right)$ and the order of queue length for each secondary node becomes $O\left(\sqrt{\frac{m}{n}}P_{off}\right)$.

- The throughput capacity and delay trade-off for CRAHNs in the regular dense primary network is given by $\Theta(\frac{WP_{off}^{\Theta(\sqrt{m})/K}}{\sqrt{n\log n}})vsO(K)$, where$1 \leq K \leq \Theta(\sqrt{m})$. Specifically, increasing the end-to-end delay results in exponential increase of the throughput capacity and vice versa.

- Under the regular sparse primary network, both the throughput capacity and delay of CRAHNs will be greatly improved. In this case, the per-node throughput capacity of CRAHN is between $\Omega(\frac{WP_{off}}{\sqrt{n\log n}})$ and $O(\frac{W}{\sqrt{n\log n}})$. The corresponding end-to-end delay is only affected by multi-hopping delay as in [10].

- Finally, we prove that the distribution of the PU locations does not impact the scalability of CRAHNs.

REFERENCES

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 201–220, Feb. 2005.

[2] I. F. Akyildiz, W. Lee, M. C. Vuran, and S. Mohanty, "Next generation/ dynamic spectrum access/congitive radio wireless networks: A survey," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 50, no. 13, pp. 2127–2159, 2006.

[3] Federal Communications Commission, *Spectrum Policy Task Force*, pp. 2–135. Cambridge, England: Rep. ET Docket 02-48, 2002.

[4] E. Hossain and V. K. Bhargava, *Cognitive Wireless Communication Networks*. Springer, 2007.

[5] I.F. Akyildiz, W. Lee, and K. Chowdhury, "CRAHNs: Cognitive Radio Ad Hoc Networks," *Ad Hoc Networks (Elsevier) Journal*, vol. 7, no. 5, pp. 810–836, 2009.

[6] M. Chiang, P. Hande, T. Lan and C.W. Tan, *Power Control in Wireless Cellular Networks*. Foundations and Trends in Networking, 2008.

[7] K. J. Ray Liu and Beibei Wang, *Cognitive Radio Networking and Security: A Game Theory Theoretical View*. Cambridge University Press, 2008.

[8] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory.*, vol. 46, no. 2, pp. 388–404, 2000.

[9] M. Grossglauser and M. D. Tse, "Mobility increases the capacity of ad hoc wireless networks," in *Proc. IEEE Inter. Conf. on Comp. Comm.*, pp. 1360–1369, 2001.

[10] A. E. Gamal, J. Mammen, B. Prabhakar and D. Shah, "Throughput-delay trade-of in wireless networks," in *Proc. IEEE Inter. Conf. on Comp. Comm.*, pp. 464–475, 2004.

[11] X. Lin, G. Sharma, R. Mazumdar, and N. B. Shroff, "Degenerate delay capacity tradeoffs in ad hoc networks with Brownian mobility," *IEEE Trans. Inf. Theory.*, vol. 52, no. 6, pp. 2777–2784, 2006.

[12] J. Herdtner and E. Chong, "Throughput-storage tradeoff in ad hoc networks," in *Proc. IEEE Inter. Conf. on Comp. Comm.*, pp. 2536–2542, 2005.

[13] B. Wang, Y. Wu and K.J. R. Liu, "Game theory for cognitive radio networks: An overview," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 54, no. 14, pp. 2537–2561, 2010.

[14] M. Maskery, V. Krishnamurthy, and Q. Zhao, "Decentralized Dynamic Spectrum Access for Cognitive Radios: Cooperative Design of a Non-cooperative Game," *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 459–469, 2009.

[15] Federal Communications Commission, *First Report and Order*. Cambridge, England: Rep. ET Docket 02-48, 2002.

[16] K.C.Chen and R. Prasad, *cognitive radio networks*. Wiley, 2009.

[17] Y. Chen, G. Yu, Z. Zhang, H. Chen, P. Qiu, "Optimal Power Allocation for Fading Channels in Cognitive Radio Networks: Ergodic capacity and outage capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 940–950, 2009.

[18] L. Qiu, Y. Yang, Y. Zhang, and S. Shenker, "On Selfish Routing in Internet-Like Environments," in *Proc. ACM SIGCOMM* , pp. 12–25, 2003.

[19] J., Shu and P., Varaiya, "Pricing Network Services," in *Proc. IEEE Conf. on Comp. Comm.* , pp. 1221–1230, 2003.

[20] A. Al Daoud, T. Alpcan, S. Agarwal, and M. Alanyali, "A Stackelberg game for pricing uplink power in wide-band cognitive radio networks," in *Proc. 47th IEEE Conf. Decision Control*, pp. 1422–1427, 2008.

[21] A. Al-Daoud, M. Alanyali, and D. Starobinski, "Secondary pricing of spectrum in cellular CDMA networks," in *Proc. IEEE DySPAN*, pp. 535–542, 2007.

[22] Y. Xing, R. Chandramouli, and C. M. Cordeiro, "Price dynamics in a competitive agile secondary spectrum access market," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 613–621, 2007.

[23] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 192–202, 2008.

[24] Y. Wu and D. H. K. Tsang, "Distributed multi-channel power allocation algorithm for spectrum sharing cognitive radio networks with QoS guarantee," in *Proc. IEEE Conf. on Comp. Comm.* , pp. 1221–1230, 2009.

[25] H. Yu, L.Gao, Z. Li, X. Wang and E. Hossain, "Pricing for uplink power control in cognitive radio networks," *IEEE Trans. Veh. Tech.*, vol. 59, no. 4, pp. 1769 – 1778, 2010.

[26] P. Zhou, W. Yuan, W. Liu and W. Cheng, "Joint power and rate control in cognitive radio networks: a game-theoretical approach," in *Proc. IEEE Inter. Conf. on Comm.*, pp. 3296– 3301, May 2008.

[27] P. Zhou, Y. Chang and J. Copeland, "Asynchronous Power Control Game with Channel Outage Constraints for Cognitive Radio Networks," in *Proc. IEEE International Conference on Communications* , pp. 1–5, 2011.

[28] X. Kang, R. Zhang, Y. C. Liang and H.K. Garg, "On Outage Capacity of Secondary Users in Fading Cognitive Radio Networks with Primary User's Outage Constraint," in *Proc. IEEE Global Telecommunications Conference*, pp. 1–5, 2009.

[29] R. Zhang, "Optimal power control over fading cognitive radio channel by exploiting primary user CSI," in *Proc. IEEE Global Telecommunications Conference*, pp. 1–5, 2008.

[30] Y. Wu, D.H.K. Tsang, "Distributed Power Allocation Algorithm for Spectrum Sharing Cognitive Radio Networks with QoS Guarantee," in *Proc. IEEE Inter. Conf. on Comp. Comm.*, pp. 981–989, 2009.

[31] H. Li and Z. Han, "Dogfight in Spectrum: Combating Primary User Emulation Attacks in Cognitive Radio Systems Part I: Known Channel Statistics," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 14–25, 2011.

[32] H. Li and Z. Han, "Dogfight in Spectrum: Combating Primary User Emulation Attacks in Cognitive Radio Systems Part II: Unknown Channel Statistics," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 26–35, 2011.

[33] M. J. Neely and E. Modiano, "Capacity and delay tradeoffs for ad hoc mobile networks," *IEEE Trans. Inf. Theory.*, vol. 51, no. 6, pp. 1917–1937, 2005.

[34] X. Lin, G. Sharma, R. Mazumdar, and N. B. Shroff, "Degenerate delaycapacity tradeoffs in ad hoc networks with Brownian mobility," *IEEE Trans. Inf. Theory.*, vol. 52, no. 6, pp. 2777–2784, 2006.

[35] S-W Jeon, N. Devroye, M. Vu, S-Y Chung and V. Tarokh, "Cognitive networks achieve throughput scaling of a homogeneous network," *IEEE Trans. Inf. Theory.*, vol. 52, no. 6, pp. 77–84, 2009.

[36] C-Y. Yin, L. Gao and S. Cui, "Scaling laws for overlaid wireless networks: a cognitive radio network vs. a primary network," in *Preprint,http://arxiv.org/pdf/0805.1209.*, 2008.

[37] V. Bhandari and N. H. Vaidya, "Connectivity and capacity of multichannel wireless networks with channel switching constraints," in *Proc. IEEE Inter. Conf. on Comp. Comm.*, pp. 785–793, 2007.

[38] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Mag.*, vol. 24, pp. 79–89, May 2007.

[39] D. Chen, S. Yin, Q. Zhang, M. Liu, and S. Li, "Mining spectrum usage data: a large-scale spectrum measurement study," in *Proc. ACM the 15th annual international conference on Mobile computing and networking*, pp. 13–24, 2009.

[40] R. D. Yates and C.-Y. Huang, "Integrated power control and base station assignment," *IEEE Trans. Veh. Tech.*, vol. 44, no. 3, pp. 638–644, 1995.

[41] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, 1995.

[42] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 291–303, 2002.

[43] J. Huang, R. Berry, and M. L. Honig, "Auction-based spectrum sharing," *ACM Mobile Netw. Appl. J.*, vol. 11, no. 3, pp. 405–418, 2006.

[44] J. Jia and Q. Zhang, "A non-cooperative power control game for secondary spectrum sharing," *ACM Mobile Netw. Appl. J.*, vol. 11, no. 3, pp. 405–418, 2006.

[45] R. Gibbons, *Game Theory for Applied Economists*. Princeton Univ. Press, 1992.

[46] D. Niyato and E. Hossain, "Integration of WiMAX and WiFi: Optimal pricing for bandwidth sharing," *IEEE Commun. Mag.*, vol. 45, pp. 140–146, May 2007.

[47] D. Kim, L. Le and E. Hossain, "Joint rate and power allocation for cognitive radios in dynamic spectrum access environment," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4710–4718, 2008.

[48] X. Kang, Y. C. Liang, A. Nallanathan, "Optimal Power Allocation for Fading Channels in Cognitive Radio Networks under Transmit and Interference Power Constraints," in *Proc. IEEE Inter. Conf. on Comm.* , pp. 3568–3572, May 2008.

[49] C. Sun, Y.D. Alemseged, H.N. Tran and H. Harada, "Transmit power control for cognitive radio over a rayleigh fading channel," *IEEE Trans. Veh. Tech.*, vol. 59, no. 4, pp. 1847–1857, 2010.

[50] X. Kang, R. Zhang, Y. C. Liang, H.K. Garg, "Optimal Power Allocation for Fading Channels in Cognitive Radio Networks under Transmit and Interference Power Constraints," in *Proc. IEEE Global Telecomm. Conf.*, pp. 1–5, 2009.

[51] C. Sun, Y.D. Alemseged, H.N. Tran and H. Harada, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency," *IEEE Commun. Mag.*, vol. 42, no. 3, pp. 8–14, 2004.

[52] G. Bansal, J. Hossain, and V. K. Bhargava, "Optimal and suboptimal power allocation schemes for ofdm-based cognitive radio systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4710–4718, 2008.

[53] P. Wang, M. Zhao, L. Xiao, S. Zhou, and J. Wang, "Power allocation in ofdm-based cognitive radio systems," in *Proc. IEEE Global Telecomm. Conf.*, pp. 4061–4065, 2007.

[54] X. Kang, H.K. Garg, Y. C. Liang, R. Zhang, "Power Allocation for OFDM-Based Cognitive Radio Systems with Hybrid Protection to Primary Users," in *Proc. IEEE Global Telecomm. Conf.*, pp. 1–6, 2009.

[55] F. E. Lapiccirella, S. Huang, X. Liu, and Z. Ding, "Feedback-based access and power control for distributed multiuser cognitive networks," in *Proc. Information Theory and Application (ITA) workshop*, pp. 1–5, 2009.

[56] S. Huang, X. Liu, and Z. Ding, "Decentralized Cognitive Radio Control based on Inference from Primary Link Control Information," *IEEE J. Sel. Areas Commun.*, vol. 1, no. 1, pp. 22–33, 2011.

[57] C. Long, Q. Zhang, B. Li, H. Yang, and X. Guan, "Non-Cooperative Power Control for Wireless Ad Hoc Networks with Repeated Games," *IEEE J. Sel. Areas Commun.*, vol. 46, no. 2, pp. 388–404, 2007.

[58] P. Zhou, Y. Chang and J. Copeland, "Learning through reinforcement for repeated power control game in cognitive radio networks," in *Proc. IEEE Global Telecommunications Conference*, pp. 1157–1162, 2010.

[59] M. Vu, N. Devroye, M. Sharif and V. Tarokh, "Scaling laws of cognitive networks," in *Proc. Inter. Conf. on Cognitive Radio Oriented Wireless Net. and Comm. (Crowncom 2007)*, pp. 2–8, May 2007.

[60] M. Chiang, *Geometric programming for communication systems*. Commun. Inf. Theory, 2005.

[61] D. Fudenberg, D. K. Levine, *The Theory of Learning in Games*. Cambridge University Press, 1998.

[62] Bush, R. and Mosteller, F., *Stochastic Models of Learning*. John Wiliey & Son, New York, 1995.

[63] V. Shah, N. B. Mandayam and D. J. Goodman, "Power control for wireless data based on utility and pricing," in *Proc. 9th IEEE Inter. Symp. on Personal, Indoor, and Mobile Radio Comm.* , pp. 1427–1432, 1998.

[64] D. J. Goodman and N. B. Mandayam, "Power control for wireless data," *IEEE Pers. Comm.*, vol. 7, no. 11, pp. 48–54, 2000.

[65] C. U. Saraydar, N. B. Mandayam and D. J. Goodman, "Pricing and power control in a multicell wireless data network," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 11, pp. 1883–1892, 2001.

[66] R. Mazumdar, L. G. Mason and C. Douligieris, "Fairness in network optimal flow control: optimality of product forms," *IEEE Trans. Commun.*, vol. 39, pp. 775–782, May 1991.

[67] L. Pavel, "An extension of duality to a game-theoretic framework," *Automatica (Journal of IFAC)*, vol. 43, no. 2, pp. 226–237, 2007.

[68] D. Knuth, *The Art of Computer Programming, Volume 1: Fundamental Algorithms*. Addison-Wesley, 1997.

[69] G.L. Stuber, *Principles of Mobile Communication*. Kiuwer Academic Press, 2001.

[70] S. Kandukuri and S. Boyd, "Optimal power control in interferencelimited fading wireless channels with outage-probability specifications," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 46–55, 2002.

[71] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1999.

[72] D. Bertsekas and J. N. Tsitsiklis, *Parellel and Distributed Computation*. Englewood Cliffs, 1989.

[73] A.J., Goldsmith and P.P., Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inf. Theory.*, vol. 42, no. 3, pp. 868–886, 1996.

[74] A.J., Goldsmith, *Wireless Communication*. Cambridge University Press, 2005.

[75] A. S. Poznyak, K. Najim, and E. Gomez, *Self-Learning Control for Finite Markov Chains*. New York: Marcel Dekker, 2000.

[76] M. Chiang, "Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 104–116, 2005.

[77] A.N., Shiryaev, *Probability*. Springer Press, 2005.

[78] P., Hall and C.C. Heyde, *Martingale Limmit Theory and Its Applications*. Academic Press, 1980.

[79] J. Nash, "Equilibrium points in n-person games," in *Proc. Nat. Acad. USA*, p. 48C49, 1950.

[80] K. Najim and A. S. Poznyak, *Learning Automata: Theory and Applications*. Springer-Verlag, 1994.

[81] H. Robbins and D. Siegmund, "A convergence theorem for non negative almost supermar-tingales and some applications," *In Optimizing Methods in Statitsics. (JS Rustagi, ed.)*, vol. 16, no. 4, pp. 233–257, 1971.

[82] K. J. Arrow, L. Hurwicz, and H. Uzawa, "Constraint qualifications in maximization problems," *Nav. Res. Logist. Q.*, vol. 8, no. 4, pp. 175–191, 1961.

[83] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave N-persons games," *Econometrica*, vol. 33, no. 6, pp. 520–534, 1965.

[84] J. P. Aubin, *Mathematical Methods of Game and Economic Theory*. Amsterdam, The Netherlands: North Holland, 1979.

[85] N. N. Vorob́ev, *Foundations of Game Theory: Noncooperative Games*. Basel: Birkhá user, 1994.

[86] Whittle P., *Optimization under Constriants*. Wiley-Interscience, New York, 1971.

[87] A. S. Poznyak and K. Najim, *Learning Automata and Stochastic Optimization*. New York: Springer-Verlag, 1997.

[88] Baba N., *New Topics in Learning Automata Theory and Applications*. Springer-Verlag, Berlin, 1984.

[89] H. S. Wang and N. Moayeri, "Finite-state Markov channel-a useful model for radio communication channels," *IEEE Trans. Veh. Tech.*, vol. 44, no. 1, pp. 163–171, 1995.

[90] P. Gupta and P. Kumar, "Critical power for asymptotic connectivity in wireless networks. Stochastic Analysis," *Control, Optimization and Applications, A Volume in Honor of W.H. Fleming*, pp. 547–566, 1998.

[91] W. Ren, Q. Zhao, and A. Swami, "Connectivity of cognitive radio networks: proximity vs. opportunity ," in *Proc. of ACM MobiCom Workshop on Cognitive Radio Networks*, 2009.

[92] W. Ren, Q. Zhao, and A. Swami, "Connectivity of heterogeneous wireless networks," *IEEE Trans. Inf. Theory.*, vol. 52, no. 8, pp. 1917–1937, 2009.

[93] K. Sriram and W. Whitt, "characterizing superposition arrial processes in packet multiplexers for voice and data," *IEEE J. Sel. Areas Commun.*, vol. 4, no. 3, pp. 883–846, 1986.

[94] J. F. Kurose and K. W. Ross, *Computer Networking, a top down approach (4th edition)*. Addision-Wesley Press, 2005.

[95] Michael Mitzenmacher, Eli Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

VITA

Pan Zhou was born in the city of Jingzhou, Hubei Province, China, in Dec, 1986. He received his B.S. degree in the *Advanced Class* of Huazhong University of Science and Technology, and a M.S. degree in Electrical Engineering from Huazhong University of Science and Technology, Wuhan, China, in 2006 and 2008, respectively. He held *honorary degree* and *merit research award* in his master study. He also received a M.S. degree in Electrical and Computer Engineering from Georgia Institute of Technology in 2010. He was once a student intern at Mobile Communications and Networking Research, NEC Laboratories America, Inc., in summer 2010. In May, 2011, he received his **Ph. D. degree** in Electrical and Computer Engineering from Georgia Institute of Technology. His research interest is majorally on wireless communications and networking problems, such as, cognitive radio, wireless scheduling, and theoretical research of network science. He is a fan of applied mathematics and operation research in MATH and ISYE department, Georgia Institute of Technology. The style of simple and mathematics related research with impressive results are his favorite.